

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 17, 2002, 08:24:35 ; Search time 1208.78 Seconds
(without alignment)
256.813 Million cell updates/sec

Title: US-09-780-113D-15
Perfect score: 23
Sequence: 1 accagcactgaagcagcactt 23

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estb1:
2: em_estb2:
3: em_estb3:
4: em_estb4:
5: em_estb5:
6: em_estb6:
7: em_estb7:
8: em_estb8:
9: em_estb9:
10: em_estb10:
11: em_estb11:
12: em_estb12:
13: em_estb13:
14: em_estb14:
15: em_estb15:
16: em_estb16:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	18.2	79.1	530 12	AA057613 HS-5097-B
2	18.2	79.1	914 10	BC165582 602343735
3	17.8	77.4	375 9	AA478086 18502 MAR
4	17.8	77.4	697 10	BI875875 1164007.Y
5	17.4	75.7	511 10	BM354761 1164007.Y
6	17.4	75.7	511 10	BM354761 1164007.Y
7	17.2	74.8	246 10	W06581 72484 MVAR4
8	17.2	74.8	455 9	AI621053 1576407.X
9	17.2	74.8	475 10	BI345371 374175 MA
10	17.2	74.8	574 12	BI622013 1007116B0
11	17.2	74.8	632 10	BI886608 2F637-1-0
12	17.2	74.8	997 12	CNS05NKN
13	17.2	74.8	1302 10	BE9652881
14	17.2	74.8	476 10	BE9652881
15	17.2	74.8	520 10	BE9652881
16	16.8	73.0	325 9	BE109045 UI-R-BS1-
17	16.8	73.0	346 9	AU052055 AU052055

C 18	16.8	73.0	386	10	BF332831
C 19	16.8	73.0	594	9	AI940637
C 20	16.8	73.0	601	9	AI940638
C 21	16.8	73.0	654	9	AM057195
C 22	16.8	73.0	896	10	BI086804
C 23	16.8	73.0	947	10	BF204564
C 24	16.8	73.0	1409	10	BF237879
C 25	16.6	72.2	297	9	BB369390
C 26	16.6	72.2	309	9	BB369390
C 27	16.6	72.2	327	9	BB369390
C 28	16.6	72.2	348	10	BF765798
C 29	16.6	72.2	435	12	AZ058336
C 30	16.6	72.2	460	10	BF036844
C 31	16.6	72.2	508	9	AA922286
C 32	16.6	72.2	522	12	AA0550432
C 33	16.6	72.2	532	10	BF768721
C 34	16.6	72.2	545	12	TA931F039
C 35	16.6	72.2	551	10	BE450880
C 36	16.6	72.2	555	12	BA403758
C 37	16.6	72.2	567	12	BB64239
C 38	16.6	72.2	574	9	AU145954
C 39	16.6	72.2	594	12	AZ093474
C 40	16.6	72.2	595	9	AM934499
C 41	16.6	72.2	622	12	AO984457
C 42	16.6	72.2	632	10	BM426935
C 43	16.6	72.2	637	12	AZ115664
C 44	16.6	72.2	648	12	AZ574567
C 45	16.6	72.2	665	12	AZ222176

ALIGNMENTS

RESULT 1
AA057613/c
DEFINITION
AA057613
ACCESSION
AA057613.1 GI:4636253
VERSION
GSS.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Fukuroyoda Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulalia; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 530)
Mahairas-G.G., Wallace, J.C., Smith, K., Swartz, J.S., Holzman, J.,
Keller, A., Shaker, R., Fulong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
Contract: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Library availability, please contact Pieter de Jong
(piter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.hsc.washington.edu
plate: 673 row: P column: 13
Seq primer: 77
Class: BAC ends
High quality sequence stop: 530.
Location/Qualifiers
1..530
/organism="Homo sapiens"

TITLE
JOURNAL
MEDLINE
COMMENT

FEATURES
source

Seq primer: -40bp from Gibco
High quality sequence slope: 479.

FEATURES

Location/Qualifiers

1..511

/organism="Glophodera rostrifrons"

/db_xref="taxon:31243"

/clone_lib="Glophodera rostrifrons 02 pGEMAL1 Smart v1"

/vec_stade="32"

/lab_host="DH10B"

/note="Vector: pGEMAL1 (Invitrogen); Site_1: BstXI;

Site_2: EcoRI; The library was donated for sequencing by

Geert Smant from Wageningen University, Laboratory of

Nematology, The Netherlands."

RESULT 8

A1621053/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

SOURCE

REFERENCE

AUTHORS

TITLE

COMMENT

JOURNAL

COMMENT

FEATURES

Location/Qualifiers

1..455

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCL_CGAP_G6"

/tissue_type="Pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7730-Pac (pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA

from the normalized library NCL_CGAP_G6 was prepared, and

as circles were made in vitro. Following BAP purification

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was FPK-amplified cDNAs from a pool

of 5,000 clones made from the same library (clones

1257095-1258631, 1469064-1470983, and 1475922-1476743).

Subtraction by Bento Soares and M. Fatima Renaldo."

FEATURES

Location/Qualifiers

1..455

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCL_CGAP_G6"

/tissue_type="Pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7730-Pac (pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA

from the normalized library NCL_CGAP_G6 was prepared, and

as circles were made in vitro. Following BAP purification

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was FPK-amplified cDNAs from a pool

of 5,000 clones made from the same library (clones

1257095-1258631, 1469064-1470983, and 1475922-1476743).

Subtraction by Bento Soares and M. Fatima Renaldo."

FEATURES

Location/Qualifiers

1..455

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCL_CGAP_G6"

/tissue_type="Pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7730-Pac (pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA

from the normalized library NCL_CGAP_G6 was prepared, and

as circles were made in vitro. Following BAP purification

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was FPK-amplified cDNAs from a pool

of 5,000 clones made from the same library (clones

1257095-1258631, 1469064-1470983, and 1475922-1476743).

Subtraction by Bento Soares and M. Fatima Renaldo."

FEATURES

Location/Qualifiers

1..455

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCL_CGAP_G6"

/tissue_type="Pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7730-Pac (pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; plasmid DNA

from the normalized library NCL_CGAP_G6 was prepared, and

as circles were made in vitro. Following BAP purification

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was FPK-amplified cDNAs from a pool

BASE COUNT 172 a 144 c 154 g 162 t
 ORIGIN
 Query Match 74.8% Score 17.2; DB 10; Length 612;
 Best Local Similarity 86.4%; Pred. No. 1.3e+03;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACGACTGACGACGACCTTT 23
 ||||| ||||| ||||| |||||
 DB 781 CCACGACGACGACGACCTGT 760

RESULT 12
 LOCUS CNS05NMK/c 997 bp DNA linear GSS 26-MAY-2000
 DEFINITION Tetradodon nigroviridis genome survey sequence T7 end of clone
 019601 of library B from Tetradodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL345488.1 GI:8239258
 VERSION AL345488
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetradodon nigroviridis.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
 Tetraodontidae; Tetradodon.
 REFERENCE 1 (bases 1 to 997)
 Koest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernol,A. and
 Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradodon nigroviridis
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 997)
 Koest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernol,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 COMMENT Human gene number estimate provided by genome wide analysis using
 Tetradodon nigroviridis DNA sequence
 http://www.genoscope.cns.fr/Tetradodon.
 FEATURES
 SOURCE Location/Qualifiers
 1..997
 /organism="Tetradodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="9601"
 /clone_lib="B"
 /note="Genome survey sequence ID : CCAR019AD001C1-end : T7"
 BASE COUNT 265 a 168 c 240 g 303 t 21 others
 ORIGIN
 Query Match 74.8% Score 17.2; DB 12; Length 997;
 Best Local Similarity 86.4%; Pred. No. 1.3e+03;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACGACTGACGACGACCTTT 23
 ||||| ||||| ||||| |||||
 DB 781 CCACGACGACGACGACCTGT 760

RESULT 13
 LOCUS BE962881/c 1302 bp mRNA linear EST 14-DEC-2000
 DEFINITION BE962881 NIH_MGC_66 Homo sapiens CDNA clone IMAGE:3856229 3',
 mRNA sequence.

ACCESSION BE962881
 VERSION BE962881.2 GI:11766244
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1302)
 NIH-MGC http://mcc.nhl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Input/Output (1999)
 COMMENT On Oct 3, 2000 this sequence version replaced g1:10575687.
 Contact: Robert Strausberg, Ph.D.
 Email: eucaps@remail.nih.gov
 Tissue procurement: DCTD/DPB
 CDNA library preparation: Life Technologies, Inc.
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LIM)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIM at:
 http://image.llnl.gov
 plate: LINC571 row: 6 column: 06
 High quality sequence stop: 177.
 Location/Qualifiers
 1..1302
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_66"
 /clone_lib="NIH_MGC_66"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pCMV-SPOK6; Site: 1; Multi-
 Site: 2; Salt: Cloned unidirectional. Primer: 5' to 3'.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

BASE COUNT 273 a 345 c 317 g 361 t 6 others
 ORIGIN
 Query Match 74.8% Score 17.2; DB 10; Length 1402;
 Best Local Similarity 86.4%; Pred. No. 1.4e+03;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACGACGACGACGACCTTT 23
 ||||| ||||| ||||| |||||
 DB 172 CCACGACGACGACGACCTTT 151

RESULT 14
 LOCUS BE652364/c 476 bp mRNA linear EST 25-APR-2001
 DEFINITION BE652364
 ACCESSION BE652364
 VERSION BE652364.1 GI:11917494
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 476)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,K.L.,
 Casas,E., Gray,J.E., White,J., Chou,Y., Fahrenkrug,S.C., Bennett,
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
 Portea,G., Holt,I., Karayancheva,S., Liang,P., Quackenbush,J. and
 Keefe,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366

Email: smithse@mail.marv.usda.gov
 Single pass sequencing. Bases called and all trimmed with phred
 v0.960904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

PCR PRIMERS

FORWARD: AGGAACAGCTATCAGCAT
 BACKWARD: GTTTCGACGTACGACG
 Plate: 65 row: 1 column: 12
 Seq primer: ATTAGGTGACACTATAG.

FEATURES

Source

1..476
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="BARC 5BOV"
 /tissue_type="pooled"
 /lab_host="DH10b"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 95 a 105 c 131 g 145 t
 ORIGIN

Query Match 73.9%; Score 17; Dh 10; Length 476;
 Best local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CACGACTGACGACGAC 19
 11111111111111111111
 Db 42H CACGACTGACGACGAC 112

RESULT 15
 HG692626/c 520 bp mRNA linear EST 02-MAY-2001
 LOCUS 342393 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION HG692626
 ACCESSION HG692626.1 GI:13944446
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Claviata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 520)
 Sonstegard, T.S., Capuco, A.V., Van Tassel, C.P., Ashwell, M.S. and
 Wells, K.D.
 Mapping of Expressed Sequence Tags from a normalized bovine mammary
 gland cDNA library
 (unpublished (2000))
 Contact: Sonstegard TS
 USDA, ARS, Beltsville Agricultural Research Center
 Bldg. 200, Rm 2A, Beltsville, MD 20705, USA
 Tel: 301 504 8416
 Fax: 301 504 8414
 Email: tads@psi.barc.usda.gov

Single pass sequencing. Bases called and all trimmed with phred
 v0.960904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.

PCR PRIMERS

FORWARD: AGGAACAGCTATCAGCAT
 BACKWARD: GTTTCGACGTACGACG
 Plate: 42 row: 1 column: 14
 Seq primer: ATTAGGTGACACTATAG.

FEATURES

Source

1..520
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="BARC 5BOV"
 /tissue_type="pooled"
 /lab_host="DH10b"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;"

Library made from pooled mRNA isolated from mammary
 tissues at eight physiological, developmental, and tissue-
 states.

BASE COUNT 116 a 135 c 140 g 129 t
 ORIGIN

Query Match 73.9%; Score 17; Dh 10; Length 476;
 Best local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CACGACTGACGACGAC 19
 11111111111111111111
 Db 293 CACGACTGACGACGAC 277

Search completed: October 17, 2002, 10:04:01
 Job time : 1217.78 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 07:01:56 ; Search time 148.222 Seconds
(without alignments)
266,418 Million cell updates/sec

Title: US-09-780-113d-15

Perfect score: 23
Sequence: 1 accacgactgacgacacccctt 23

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : N_Geneseq_032802:**

1: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgcdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	16.6	72.2	574	22	AAH09581	Human cDNA clone (
2	16.6	72.2	736	20	AAV69441	Banana fruit ripen
3	16.6	72.2	736	20	AAV69443	Banana fruit ripen
4	16.6	72.2	787	20	AAV69442	Banana fruit ripen
5	16.6	72.2	809	20	AAV69440	Banana fruit ripen
6	16.6	72.2	1181	20	AAV27502	Banana ethylene-ri
7	16.6	72.2	1181	22	AAV09902	DNA encoding ethyl
8	16.6	72.2	1693	22	AAH13873	Human cDNA sequen
9	16.6	72.2	2319	23	AAV91632	DNA encoding novel

C 10	16.6	72.2	10079	16	AA088057	Human EC-SOD gene.
C 11	16.6	72.2	10079	18	AAV92317	Human EC-SOD gene.
C 12	16.6	72.2	10079	21	AAV28294	Human superoxide d
C 13	16.6	72.2	14872	19	AAV52205	Streptococcus pneu
C 14	16.2	70.4	744	23	AB127153	Drosophila melanog
15	16.2	70.4	2510	23	AB127146	Drosophila melanog
16	16.2	70.4	2990	23	AB127152	Drosophila melanog
C 17	15.8	68.7	380	20	AAV88222	EST clone DX295.
C 18	15.8	68.7	828	22	AAV65951	Human immune/hema
C 19	15.8	68.7	828	22	AAV65952	Human immune/hema
C 20	15.8	68.7	932	22	AAV58150	Human polynucleoli
C 21	15.8	68.7	972	23	AAV58527	DNA encoding novel
C 22	15.8	68.7	1552	23	AB105193	Drosophila melanog
C 23	15.8	68.7	2643	22	ABH33583	Human colon cancer
C 24	15.8	68.7	3608	23	ABH05192	Drosophila melanog
C 25	15.8	68.7	5684	21	AAV59953	Human secreted pro
C 26	15.8	68.7	10126	22	AAV99260	Human excretory re
C 27	15.8	68.7	10126	22	AAV63610	Human kidney relat
C 28	15.6	67.8	1227	19	AAV96124	Solidoline quicosey
C 29	15.6	67.8	519	21	AAV54157	Neisseria meningit
C 30	15.6	67.8	519	21	AAV54158	Neisseria meningit
C 31	15.6	67.8	573	23	AB127647	Drosophila melanog
C 32	15.6	67.8	608	23	AAV91030	DNA encoding novel
C 33	15.6	67.8	611	21	AAV98099	Human colon cancer
C 34	15.6	67.8	663	22	AAV93954	Human foetal cDNA,
C 35	15.6	67.8	686	22	AAV21884	Human collagen gen
C 36	15.6	67.8	855	22	AAH04495	Human cDNA clone (
C 37	15.6	67.8	919	22	AAH48496	Japanese medaka fe
C 38	15.6	67.8	1398	22	AAH17346	Human cDNA sequenc
C 39	15.6	67.8	2397	23	AB126239	Drosophila melanog
C 40	15.6	67.8	2573	23	AB127646	Drosophila melanog
C 41	15.6	67.8	2820	24	ABH04069	Human semaphorin G
C 42	15.6	67.8	285	24	ABH04068	Human semaphorin G
C 43	15.6	67.8	3105	24	ABH04064	Human semaphorin G
C 44	15.6	67.8	3150	24	ABH04062	Human semaphorin G
C 45	15.6	67.8	3237	24	ABH04065	Human semaphorin G

ALIGNMENTS

RESULT 1	AAH09581/c
ID	AAH09581 standard; cDNA; 574 BP.
AC	AAH09581;
XX	
XX	
DT	26-JUN-2001 (first entry)
XX	
DE	Human cDNA clone (3'-primer) SEQ ID NO:6416.
XX	
KW	Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
PM	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PF	28-JUL-2000; 2000BP-0116126.
XX	
PR	29-JUL-1999; 99JP-0248036.
PR	27-AUG-1999; 99JP-0300253.
PR	11-JAN-2000; 2000JP-0118776.
PR	02-MAY-2000; 2000JP-0183767.
PR	09-JUN-2000; 2000JP-0241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	
DR	WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 3; SEQ ID 6416: 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo- or primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 574 BP; 118 A; 144 C; 162 G; 143 T; 7 other;
XX
Query Match 72.2%; Score 16.6; DB 22; Length 574;
Best Local Similarity 82.6%; Pred. No. 79;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCGACGACTGAGCAGCAGCCTT 23
||||| | | | | | | | | | |
DB 312 ACCGACGCGGATCAGCAGCCTCT 290
XX
RESULT 2
AAV69441
ID AAV69441 standard; cDNA; 736 BP.
XX
AC AAV69441;
XX
DE 15-MAR-1999 (first entry)
XX
DE Banana fruit ripening-related clone U-13 cDNA.
XX
KM Fruit ripening; banana; modulator; tissue senescence; crop; plant;
KM triploid; plant breeding; ss.
XX
OS Musa acuminata.
XX
PN WO9853085-A1.
XX
PD 26-NOV-1998.
XX
PF 05-MAY-1998; 98WO-GB01297.
XX
PR 20-MAY-1997; 97GB-0010370.
XX
PA (ZENE) ZENECA LTD.
XX
PI Bird CR, Medina-Suarez RDJ, Seymour GB;
XX
DR WPI; 1999-059745/05.
PT New method of modulating fruit ripening or tissue senescence
PT characteristics of Musa plants - by introducing DNA sequences,

PT useful in processes for modifying plant/fruit ripening
PT characteristics
XX
PS Claim 1; Page 19; 78pp; English.
XX
CC AAV69440-V69512 are cDNA sequences which are used in a method to
CC modulate the fruit ripening or tissue senescence characteristics of Musa
CC acuminata (banana) plants. The method provides a recombinant way of
CC modulating ripening/senescence characteristics of bananas, which are a
CC globally important crop. In particular, the method can modulate such
CC characteristics in the desert banana, which due to its triploid nature
CC is unaffected by conventional plant breeding techniques.
XX
SQ Sequence 736 BP; 186 A; 190 C; 196 G; 146 T; 18 other;
XX
Query Match 72.2%; Score 16.6; DB 20; Length 736;
Best Local Similarity 82.6%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCGACGCTGAGCAGCAGCCTT 23
||||| | | | | | | | | | |
DB 281 ACCTGACTGAGCAGCAGCCTT 303
XX
RESULT 3
AAV69443
ID AAV69443 standard; cDNA; 736 BP.
XX
AC AAV69443;
XX
DE 15-MAR-1999 (first entry)
XX
DE Banana fruit ripening-related clone U-105 cDNA.
XX
KM Fruit ripening; banana; modulator; tissue senescence; crop; plant;
KM triploid; plant breeding; ss.
XX
OS Musa acuminata.
XX
PN WO9853085-A1.
XX
PD 26-NOV-1998.
XX
PF 05-MAY-1998; 98WO-GB01297.
XX
PR 20-MAY-1997; 97GB-0010370.
XX
PA (ZENE) ZENECA LTD.
XX
PI Bird CR, Medina-Suarez RDJ, Seymour GB;
XX
DR WPI; 1999-059745/05.
PT New method of modulating fruit ripening or tissue senescence
PT characteristics of Musa plants - by introducing DNA sequences,
PT useful in processes for modifying plant/fruit ripening
PT characteristics
XX
PS Claim 1; Page 20; 78pp; English.
XX
CC AAV69440-V69512 are cDNA sequences which are used in a method to
CC modulate the fruit ripening or tissue senescence characteristics of Musa
CC acuminata (banana) plants. The method provides a recombinant way of
CC modulating ripening/senescence characteristics of bananas, which are a
CC globally important crop. In particular, the method can modulate such
CC characteristics in the desert banana, which due to its triploid nature
CC is unaffected by conventional plant breeding techniques.
XX
SQ Sequence 736 BP; 189 A; 193 C; 202 G; 138 T; 14 other;
XX
Query Match 72.2%; Score 16.6; DB 20; Length 736;
Best Local Similarity 82.6%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCAGACTGACGACCTTT 23
 DB 284 ACCTGACTGGGAAGCACCCTT 306

RESULT 4

AAV69442
 ID AAV69442 standard; cDNA: 787 BP.

AC AAV69442;

DT 15-MAR-1999 (first entry)

DE Banana fruit ripening-related clone U-23 cDNA.

KW Fruit ripening; banana; modulator; tissue senescence; crop; plant;

KW triploid; plant breeding; ss.

OS Musa acuminata.

PN WO9853085-A1.

PD 26-NOV-1998.

PF 05-MAY-1998; 98WO-GB01297.

PR 20-MAY-1997; 97GB-0010370.

PA (ZENE) ZENECA LTD.

PI Bird CR, Medina-Suarez RDU, Seymour GB;

DR WPI: 1999-059745/05.

PT New method of modulating fruit ripening or tissue senescence

PT characteristics of Musa plants - by introducing DNA sequences,

PT useful in processes for modifying plant/fruit ripening

PT characteristics

PS Claim 1; Page 19-20; 78pp; English.

CC AAV69440-V69512 are cDNA sequences which are used in a method to

CC modulate the fruit ripening or tissue senescence characteristics of Musa

CC acuminata (banana) plants. The method provides a recombinant way of

CC modulating ripening/senescence characteristics of bananas, which are a

CC globally important crop. In particular, the method can modulate such

CC characteristics in the desert banana, which due to its triploid nature

CC is unaffected by conventional plant breeding techniques.

CC

SO Sequence 787 BP; 209 A; 206 C; 216 G; 150 T; 6 other;

Query Match 72.2%; Score 16.6; DB 20; Length 787;

Best Local Similarity 82.6%; Pred. No. 82;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCAGACTGACGACCTTT 23

DB 284 ACCTGACTGGGAAGCACCCTT 306

RESULT 5

AAV69440
 ID AAV69440 standard; cDNA: 809 BP.

AC AAV69440;

DT 15-MAR-1999 (first entry)

DE Banana fruit ripening-related clone U-7 cDNA.

KW Fruit ripening; banana; modulator; tissue senescence; crop; plant;

KW triploid; plant breeding; ss.

XX OS Musa acuminata.

XX PN WO9853085-A1.

XX PD 26-NOV-1998.

XX PF 05-MAY-1998; 98WO-GB01297.

XX PR 20-MAY-1997; 97GB-0010370.

XX PA (ZENE) ZENECA LTD.

XX PI Bird CR, Medina-Suarez RDU, Seymour GB;

XX DR WPI: 1999-059745/05.

XX PT New method of modulating fruit ripening or tissue senescence

XX PT characteristics of Musa plants - by introducing DNA sequences,

XX PT useful in processes for modifying plant/fruit ripening

XX PT characteristics

XX PS Claim 1; Page 18; 78pp; English.

XX CC AAV69440-V69512 are cDNA sequences which are used in a method to

XX CC modulate the fruit ripening or tissue senescence characteristics of Musa

XX CC acuminata (banana) plants. The method provides a recombinant way of

XX CC modulating ripening/senescence characteristics of bananas, which are a

XX CC globally important crop. In particular, the method can modulate such

XX CC characteristics in the desert banana, which due to its triploid nature

XX CC is unaffected by conventional plant breeding techniques.

XX

SO Sequence 809 BP; 210 A; 214 C; 217 G; 153 T; 15 other;

Query Match 72.2%; Score 16.6; DB 20; Length 809;

Best Local Similarity 82.6%; Pred. No. 82;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCAGACTGACGACCTTT 23

DB 277 ACCTGACTGGGAAGCACCCTT 299

RESULT 6

AAV27502

ID AAV27502 standard; cDNA: 1181 BP.

XX AX27502;

XX 26-MAY-1999 (first entry)

XX Banana ethylene-forming enzyme cDNA sequence.

XX 1-aminocyclopropane-1-carboxylic acid synthase; ACS; ERF; banana;

XX ethylene forming enzyme; ethylene biosynthesis; plant; fruit ripening;

XX transgenic; enzyme; inhibition; flavour; texture; ss.

XX Musa sp.

XX US5886164-A.

XX 23-MAR-1999.

XX 15-APR-1996; 96US-0632598.

XX 15-APR-1996; 96US-0632598.

XX (ZENE) ZENECA LTD.

XX Bird CR, Fletcher JD;

XX WPI: 1999-228611/19.

Pt	1-aminocyclopropane-1-carboxylic acid synthase or ethylene-forming enzyme
Pt	-
Ps	Claim 2; Column 15: 23pp; English.
Xx	
Cc	The sequence represents the coding sequence of ethylene-forming enzyme (EFE) from banana, used in the method of the invention. The method involves modifying the level of ethylene biosynthesis in a plant of the genus Musa by inserting into the genome of the plant a DNA sequence (1) encoding a banana 1-aminocyclopropane-1-carboxylic acid synthase (ACS) or an ethylene-forming enzyme (EFE), where (1) is in sense or antisense configuration, and modifies the level of activity of ACS or EFE. This retards the rate of ripening in banana fruits which reduces the rate of deterioration of banana fruit after harvest. As a result, fruit may be harvested when they have reached partial or full ripeness and still have the robustness to withstand handling and transport to reach the consumer in good condition. In this way high quality ripe fruit can be made available to the consumer with reduced requirement for post-harvest treatment. High quality fruit will have improved flavour and texture.
Cc	High quality fruit can be produced consistently over a wide harvest period, and such fruit can be held in store for long periods and ripened to optimal quality by the supply of exogenous ethylene.
Cc	
Sq	Sequence 1181 BP; 319 A; 294 C; 318 G; 250 T; 0 other;
Qy	Query Match
Db	Best Local Similarity 72.2%; Score 16.6; DB 22; Length 1181; Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0.
Oy	1 ACCACGACTGAGCAGCACCTT 23
DB	289 ACCTCGACTGGGAAGCACCTT 111
AAH13873	
AAH13873	Standard; cDNA; 1693 BP.
AAH13873:	
26-JUN-2001	(first entry)
Human cDNA sequence SEQ ID NO:10870.	
Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.	
Homo sapiens.	
EPI074617-A2.	
07-FEB-2001.	
28-JUL-2000; 2000EP-0116126.	
26-AUG-1999; 98JP-0248036.	
27-AUG-1999; 98JP-0300253.	
11-JAN-2000; 2000JP-0118776.	
02-MAY-2000; 2000JP-0183767.	
09-JUN-2000; 2000JP-0241899.	
(HELI-) HELIX RES INST.	
Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T; WPI; 2001-318749/34.	
Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -	
Claim 8; SEQ ID 10870; 2537bp + CD ROM; English.	


```

FT      /label= sis responsive element
FT      162..168
FT      /tag= j
FT      /label= Ap1 binding site/TPA responsive element
FT      171..179
FT      /tag= k
FT      /label= SV40 enhancer region Ap4
XX
XX      W09510185-A.
XX
XX      20-APR-1995.
XX
XX      13-OCT-1994; 94MO-US11558.
XX
XX      15-OCT-1993; 93US-0136207.
XX      20-SEP-1994; 94US-0136207.
XX
XX      (UYAL-) UNIV ALABAMA.
XX      (UYDU-) UNIV DUKE.
XX
XX      Crapo JD, Day BJ, Koiz RJ, Freeman BA, Fridovich I;
XX      Oury T, Foltz J;
XX
XX      WPI; 1995-161483/21.
XX      P-PSDB; AAR72413.
XX
XX      Super-oxide dismutase mimetic(s) comprising a nitrogen-containing
XX      macrocyclic moiety - useful in treatment of inflammation,
XX      musculetory disorders, atherosclerosis or gout.
XX
XX      Claim 33; Figure 24; 136pp; English.
XX
XX      Clone 7 from a human adult leukocyte genomic library was analysed.
XX      The analysis revealed that human EC-SOD contains three exons and
XX      two introns. The intron-exon boundaries are shown in AA088062-088065.
XX      AA088057 shows the entire sequence for the human EC-SOD gene. Exon 3
XX      contains the entire uninterrupted coding region. An obvious TATA
XX      box cannot be discerned for the EC-SOD gene. Two CAT box elements
XX      were identified. One is in the reverse orientation and located
XX      about 20bp upstream of the first exon. Transcriptional factor
XX      database searching of the 5' nontranslated region and first exon
XX      revealed several potential regulatory elements (see Fig).
XX
XX      Sequence 10079 BP: 2482 A; 2613 C; 2407 G; 2577 T; 0 other;
XX
XX      Query Match 72.2%; Score 16.6; DB 16; Length 10079;
XX      Best Local Similarity 82.6%; Pred. No. 1.1e+02;
XX      Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY 1 ACCACGACGTGAGCAGCACCCTTT 23
XX      1238 ACCTCACTTAGCACCACCCTTT 1216
XX
XX      RESULT 11
XX      AAT92317/c
XX      ID AAT92317 standard; cDNA; 10079 BP.
XX
XX      AC AAT92317;
XX
XX      DT 06-FEB-1998 (first entry)
XX
XX      DE Human EC-SOD gene.
XX
XX      KW oxidant scavenger; extracellular matrix targeting moiety;
XX      cell surface matrix; nitrogen containing macrocyclic moiety;
XX      inflammatory condition; aberrant smooth muscle function;
XX      ischaemia reperfusion injury; myocardial infarction; stroke;
XX      acute head trauma; organ reperfusion; bowel ischaemia;
XX      pulmonary infarction; glaucoma; EC-SOD; superoxide dismutase;
XX      tetrameric glycosylated copper and zinc containing enzyme; ss.
XX      Homo sapiens.

```

```

XX      XX
XX      key location/Qualifiers
XX      CDS 5085..5807
XX      /tag= a
XX      /product= EC-SOD
XX
XX      W09640223-A1.
XX
XX      19-DEC-1996.
XX
XX      07-JUN-1996; 96MO-US10497.
XX
XX      11-MAR-1996; 96US-0613418.
XX      07-JUN-1995; 95US-0476066.
XX
XX      (TROV/) TROVA M P.
XX      (UYAL-) UNIV ALABAMA.
XX      (UYDU-) UNIV DUKE.
XX
XX      Balnic-Haberle I, Crapo JD, Day BJ, Foltz RJ, Freeman BA;
XX      Fridovich I, Oury T, Trova MP;
XX
XX      WPI; 1997-077220/07.
XX      P-PSDB; AAM32599.
XX
XX      New porphyrin-type oxidant scavengers - used for protecting against
XX      oxidants and for modulating biological processes involving oxidants.
XX
XX      Claim 24; Fig 24; 195pp; English.
XX
XX      CC Oxidant scavengers have been developed comprising a nitrogen containing
XX      macrocyclic moiety and a cell surface or extracellular matrix targeting
XX      moiety, or their salts. The present sequence encodes the human EC-SOD,
XX      a tetrameric glycosylated copper and zinc containing superoxide
XX      dismutase enzyme. The oxidant scavengers can be used for protecting
XX      against the deleterious effects of oxidants and for modulating
XX      biological processes involving oxidants. They can be used for e.g.
XX      treating inflammatory conditions, treating disorders resulting from
XX      aberrant smooth muscle function or to protect against ischaemia
XX      reperfusion injuries associated with myocardial infarction, stroke,
XX      acute head trauma, organ reperfusion following transplantation, bowel
XX      ischaemia, pulmonary infarction, surgical occlusion of blood flow, and
XX      soft tissue injury. They can further be used to protect against damage
XX      to the eye due to sunlight (and to the skin) as well as glaucoma, and
XX      macular degeneration of the eye. Diseases of the bone are also amenable
XX      to treatment with the compounds, and connective tissue disorders
XX      associated with defects in collagen synthesis or degradation can be
XX      treated with the compounds. In the oxidant scavengers, substituents can
XX      be selected to render them resistant to degradation by haemoxigenase and
XX      also so that they do not interfere with normal porphyrin metabolism, can
XX      pass through cell membranes and bind to cell surface or extracellular
XX      matrix elements.
XX
XX      Sequence 10079 BP; 2481 A; 2612 C; 2409 G; 2577 T; 0 other;
XX
XX      Query Match 72.2%; Score 16.6; DB 18; Length 10079;
XX      Best Local Similarity 82.6%; Pred. No. 1.1e+02;
XX      Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX      QY 1 ACCACGACGTGAGCAGCACCCTTT 23
XX      1238 ACCTCACTTAGCACCACCCTTT 1216
XX
XX      RESULT 12
XX      AAA28294/c
XX      ID AAA28294 standard; DNA; 10079 BP.
XX
XX      AC AAA28294;
XX
XX      DT 12-FEB-2001 (first entry)
XX
XX      DE Human superoxide dismutase (EC-SOD) gene.

```

XX	Porphin; metal complex oxidant scavenger; inhibitor; EC-SOD; gout;
KW	superoxide dismutase; myocardial infarction; stroke; acute head trauma;
KM	organ reperfusion; bowel ischemia; pulmonary infarction; glaucoma; AIDS;
KM	skeletal muscle reperfusion injury; central nervous system disease; AIDS;
KW	dementia; stroke; amyotrophic lateral sclerosis; Parkinson's disease;
KW	Huntington's disease; neurological disorder; arthritis; hypertension;
KW	atherosclerosis; oedema; septic shock; pulmonary hypertension; asthma;
KW	impotence; infertility; endometriosis; diabetes; pneumonia; human;
KM	cystic fibrosis; sinusitis; autoimmune disease; ds.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 5085..5807
FT	/tag= a
FT	/product= "EC-SOD"
FT	/note= "Superoxide dismutase"
FT	sig_peptide 5085..5138
FT	/tag= b
FT	mat_peptide 5139..5804
FT	/tag= c
XX	
FN	US6127356-A.
XX	
PD	03-OCT-2000.
XX	
PE	07-JUN-1996; 96US-0663028.
XX	
PR	15-OCT-1993; 93US-0136207.
PR	13-OCT-1994; 94US-0322766.
PR	07-JUN-1995; 95US-0476866.
PR	11-MAR-1996; 96US-0613418.
XX	
PA	(UNDU-) UNIV DUKE.
XX	
P1	Crapo JD, Fridovich I, Cury T, Foiz RJ, Trova MF, Freeman BA:
P1	Battinic-Haberle I, Day BJ:
XX	
DR	WP1: 2000-664150/64.
XX	
PS	P-PSDB: AAY94782.
PT	New metal complexes of methine substituted porphines useful as
XX	catalytic oxygen scavengers -
XX	
PS	Disclosure: Fig 24; 97pp; English.
CC	This invention relates to porphines and their metal complex oxidant
CC	scavengers, where the metal is manganese, copper or iron. The porphines
CC	exhibit caditant; cerebroprotective; vasotropic; opthalmological;
CC	antiParkinsonian; nootropic; anticonvulsant; cytostatic; gynecological;
CC	antiarteriosclerotic; antiinflammatory; antibacterial; immunosuppressive;
CC	hypotensive; antidiabetic; antigout; antiaesthetic; and virucide
CC	activity. The porphines act as superoxide dismutase (SOD) inhibitors and
CC	are used as catalytic scavengers of reactive oxygen species to protect
CC	against ischemia reperfusion injuries associated with myocardial
CC	infarction, stroke, acute head trauma, organ reperfusion following
CC	transplantation, bowel ischemia, pulmonary infarction, surgical
CC	occlusion of blood flow, soft tissue injury, skeletal muscle reperfusion
CC	injuries, glaucoma, macular degeneration of the eye, diseases of the
CC	bones, to increase the limited storage viability of transplanted hearts,
CC	kidneys, skin and other organs and tissues. The compounds are also useful
CC	in the treatment of diseases of the central nervous system (including
CC	AIDS dementia, stroke, amyotrophic lateral sclerosis), Parkinson's
CC	disease, Huntington's disease, disease of the musculature, cardiac
CC	fatigue of congestive heart failure, muscle weakness syndrome associated
CC	with myopathies, neurological disorders, arthritis, systemic
CC	hypertension, arteriosclerosis, oedema, septic shock, pulmonary
CC	contractions, microbial infections, gout, type II diabetes mellitus,
CC	inflammation of the lungs, asthma, pneumonia, cystic fibrosis, chronic
CC	sinusitis and autoimmune disease. The present sequence represents the
CC	human EC-SOD gene. EC-SOD is a tetrameric glycosylated copper and zinc

CC	containing enzyme. The enzyme is used to illustrate the activity of the
CC	porphyrins of the invention.
XX	
SO	Sequence 10079 BP; 2482 A; 2612 C; 2408 G; 2577 T; 0 other;
QY	
	Query Match 72.2%; Score 16.6; DB 21; Length 10079;
	Best Local Similarity 82.6%; Pred.No. 1.1e+02;
	Matches 19; Conservative 0; Mismatches 4; Indels 0; Caps 0;
Db	
	1 ACACGACTGAGCACGCACCTT 23
	1238 ACCCTAACCTPACGACCACCTTT 1216
RESULT 13	
ID	AAV52205/c
XX	AAV52205 standard; DNA; 14872 BP.
AC	
XX	AAV52205;
DT	23-OCT-1998 (first entry)
DE	
XX	Streptococcus pneumoniae genome fragment SEQ ID NO:72.
XX	
KW	Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
RW	computer readable medium; vaccine; pharmaceutical composition; ds.
XX	
CS	Streptococcus pneumoniae.
XX	
FN	WO9818931-A2.
XX	
PD	07-MAY-1998.
XX	
PF	30-OCT-1997; 97WO-USJ9588.
XX	
PR	31-OCT-1996; 96US-0029960.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI	Kunsch CA, Rosen CA;
XX	
DR	WPI; 1998-272225/24.
XX	
PT	Computer-readable medium with recorded Streptococcus pneumoniae
PT	polynucleotide sequences - useful in diagnostic kits and assays, and
PT	pharmaceutical compositions and vaccines for Streptococcus
PT	pneumoniae
XX	
PS	Claim 1: Page 593-602; 1409pp; English.
XX	
CC	The present invention describes a computer readable medium which has
CC	the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
CC	recorded on it, or a representative fragment or a sequence at least 95%
CC	identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
CC	SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
CC	Streptococcus pneumoniae. The present invention also describes an
CC	isolated nucleic acid molecule encoding a homologue of any of the
CC	fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC	nucleic acid molecule is produced by a process comprising: (a) screening
CC	a genomic DNA library using as a probe a target sequence defined by any
CC	of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC	library which contain sequences that hybridize to the target sequence and
CC	isolating the nucleic acid molecules from the members; or (b) isolating
CC	mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC	molecules whose nucleotide sequence is homologous to amplification
CC	primers derived from the fragment of the S. pneumoniae genome to prime
CC	the amplification and isolating the amplified sequences. The computer
CC	readable medium can be used in a computer-based system for identifying
CC	fragments of the S. pneumoniae genome of commercial importance, or
CC	expression modulating fragments of the S. pneumoniae genome. Products
CC	from the present invention can be used in diagnosis kits and assays, and
CC	pharmaceutical compositions and vaccines for S. pneumoniae.

XX Sequence 14872 BP; 4340 A; 2874 C; 3210 G; 4448 T; 0 other;
SO

Query Match 72.2%; Score 16.6; DB 19; Length 14872;
Best Local Similarity 82.6%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCAGCAGCTGACGACGACCTT 23
DB 11432 ACCAGGAGTACGACGACGACCTTT 11410

RESULT 14
ABL27153
ID ABL27153 standard; DNA: 744 BP.
XX
AC ABL27153;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32932.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
XX

New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX

Claim 1; SEQ ID NO 32932; 21pp + Sequence Listing; English.
PS
XX

The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL1840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX

Sequence 744 BP; 215 A; 205 C; 202 G; 122 T; 0 other;
SQ

Query Match 70.4%; Score 16.2; DB 23; Length 744;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ACCAGCAGCTGACGACGACCT 21
DB 682 ACCAGGAGTACGACGACGACCT 702

RESULT 15
ABL27146

ID ABL27146 standard; DNA: 2510 BP.
XX
XX ABL27146;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 32911.
DE
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
XX

New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX

Claim 1; SEQ ID NO 32911; 21pp + Sequence Listing; English.
PS
XX

The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL1840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX

Sequence 2510 BP; 680 A; 581 C; 586 G; 663 T; 0 other;
SQ

Query Match 70.4%; Score 16.2; DB 23; Length 2510;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ACCAGCAGCTGACGACGACCT 21
DB 848 ACCAGGAGTACGACGACGACCT 868

Search completed: October 17, 2002, 08:50:51
Job time : 160.222 secs

```

FT      misc_RNA      /product= 28S_rRNA
FT      5583..9396
FT      /*tag= f
FT      /function= spacer
FT      misc_RNA      7338..8291
FT      /*tag= g
FT      /function= probe
FT      /note= "claimed"
FT      misc_RNA      9397..9558
FT      /*tag= h
FT      /product= 18S_rRNA
XX
XX
PN      JP07059577-A.
XX
PD      07-MAR-1995.
XX
XX      23-AUG-1993; 93JP-0227806.
XX
XX      23-AUG-1993; 93JP-0227806.
XX
XX      (KIRI ) KIRIN BEVERAGE KK.
XX
XX      WPI: 1995-135901/18.
XX
XX      Valencia orange ribosomal RNA gene - also probes and primers
XX      derived from citrus fruit RNA gene spacer regions, useful for
XX      RFLP analysis of citrus fruit
XX
XX      Claim 1; Page 7-10; 18pp; Japanese.
XX
XX      The ribosomal RNA gene has been isolated from Valencia oranges and
XX      sequenced (see AA088228). Primers based on sequences in the rRNA gene
XX      (see AA088228-088230 which are derived from nucleotides 1556-1575 and
XX      2257-2276 of AA088228, respectively) can be used to amplify regions of
XX      rRNA genes from various different types of citrus fruits. RFLP
XX      analysis and DNA fingerprinting of the amplified fragments allows
XX      different species to be distinguished.
XX
XX      Sequence 9558 BP; 1853 A; 2675 C; 2996 G; 2034 T; 0 other;
SO
Query Match      80.9%; Score 17.8; DB 16; Length 9558;
Best Local Similarity 90.5%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 CCGCTTCACTGCCGCTTACTA 21
      111111111111111111111111
DB      2340 CCGTTCGCTCCGCTTACTA 2320

RESULT 15
AAF22281/c
ID      AAF22281 standard; DNA: 59590 BP.
XX
XX      AAF22281;
XX
XX      20-MAR-2001 (first entry)
XX
XX      RAC containing repeats from centromeres 1-4 #4.
XX
XX      Centromere; michrosome; vector; ds.
XX
XX      Arabidopsis thaliana.
XX
XX      W0200055325-A2.
XX
XX      21-SEP-2000.
XX
XX      17-MAR-2000; 2000WO-US07392.
XX
XX      18-MAR-1999; 99US-0125219.
XX      01-APR-1999; 99US-0127409.
XX      18-MAY-1999; 99US-0134770.
XX      13-SEP-1999; 99US-0153584.

```

```

PR      17-SEP-1999; 99US-0154603.
XX
XX      (UYCH-) UNIV CHICAGO.
XX
XX      Preuss D, Copenhaver G, Keith K;
XX
XX      WPI: 2000-587529/55.
XX
XX      Recombinant DNA construct comprising a plant centromere, useful for
XX      producing stably inherited michrosomes which can serve as vectors for
XX      the construction of transgenic plant and animal cells
XX
XX      Claim 102; Page 351-364; 1449pp; English.
XX
XX      The present invention relates to a recombinant DNA construct of a plant
XX      (Arabidopsis thaliana) centromere. The constructs are useful for
XX      producing stably inherited michrosomes which can serve as vectors for
XX      the construction of transgenic plant and animal cells expressing
XX      selected proteins such as hormones, enzymes, interleukins, clotting
XX      factors, cytokines, antibodies, and growth factors.
XX
XX      Sequence 59590 BP; 17614 A; 12031 C; 13575 G; 16118 T; 252 other;
SO
Query Match      80.9%; Score 17.8; DB 21; Length 59590;
Best Local Similarity 90.5%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 CCGCTTCACTGCCGCTTACTA 21
      111111111111111111111111
DB      45271 CCGTTCGCTCCGCTTACTA 45251

```

Search completed: October 17, 2002, 08:50:39
 Job time : 159.778 secs


```

XX Arabidopsis thaliana; rRNA gene; ribosomal DNA; intergenic region;
KW DNA construct; plant; multiple SalI repeat; stability; copy number;
KW transgene; antibody; insecticidal protein; Bt toxin; ss.
XX Synthetic.
XX
XX WO9813505-A1.
XX
XX 02-APR-1998.
XX
XX 23-SEP-1997; 97WO-EP05217.
XX
XX 24-SEP-1996; 96AT-0001695.
XX
XX (PLB2 ) PLANT GENETIC SYSTEMS NV.
XX
XX Bachmaier A, Schweizer D;
XX
XX WPI; 1998-286413/25.
XX
XX DNA construct containing plant intergenic ribosomal DNA fragment -
XX containing multiple SalI repeats; increases stability and copy
XX number of a transgene(s) in plants
XX
XX Disclosure; Page 11; 48pp; English.
XX
XX A DNA construct has been developed which comprises the following
XX operably linked DNA fragments: (a) a ribosomal DNA sequence, preferably
XX derived from a plant; (b) an expressible (especially plant-expressible)
XX promoter region; (c) a heterologous coding region; and (d) a
XX transcription termination and polyadenylation region which preferably is
XX active in plant cells. The present sequence represents an oligonucleotide
XX from the present invention. Transformed plant cells and transgenic plants
XX comprising the DNA constructs are used to produce a desired protein at a
XX high yield, e.g. antibodies, insecticidal proteins (e.g. a Bt toxin), a
XX protein useful in the food industry. Alternatively the constructs can be
XX used to produce an antisense RNA or ribozyme. The use of intergenic
XX ribosomal DNA enhances stability and the copy number or expression of
XX transgenes in a plant.
XX
XX Sequence 24 BP; 3 A; 9 C; 6 G; 6 T; 0 other;
XX
XX Query Match 80.9%; Score 17.8; DB 19; Length 24;
XX Best Local Similarity 90.5%; Pred. No. 12;
XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CCGCTTCAGTCGCCGTACTA 21
XX ||||| ||||| |||||
XX Db 2 CCGCTTCAGTCGCCGTACTA 22
XX
XX RESULT 13
XX AAV72868/C
XX ID AAV72868 standard; DNA; 3499 BP.
XX
XX AAV72868;
XX
XX 29-MAR-1999 (first entry)
XX
XX Neospora caninum large subunit ribosomal DNA.
XX
XX large subunit ribosomal DNA; LSU-rDNA; detection; diagnosis; assay;
XX ss.
XX
XX Neospora caninum (NC-Liverpool strain).
XX
XX WO9853074-A1.
XX
XX 26-NOV-1998.
XX
XX 19-MAY-1998; 98WO-AU00367.
XX

```

```

PR 20-MAY-1997; 97AU-0006903.
XX
XX (INSE-) INSEARCHE LTD.
XX
XX Amoyal GG, Ellis JT, Rice CL;
XX
XX WPI; 1999-045313/04.
XX
XX Large subunit ribosomal DNA of Neospora species, especially N.
XX caninum - useful to derive nucleic acid sequences for isolation of
XX Neospora species by PCR amplification, e.g. to diagnose Neospora in
XX clinical samples
XX
XX Claim 1; Fig 2; 40pp; English.
XX
XX This nucleotide sequence encodes the large subunit (LSU) ribosomal
XX DNA (rDNA) of Neospora caninum. It is a compilation of overlapping
XX fragments obtained by PCR amplification of N. caninum (NC-Liverpool
XX strain) tachyzoite genomic DNA. N. caninum is a cyst-forming
XX coccidian parasite that causes neuromuscular disease in canines and
XX is a significant cause of abortion and neonatal mortality in
XX livestock. It is recognised as being closely related to Toxoplasma
XX gondii. Comparison of LSU rDNA of N. caninum with a consensus
XX sequence derived for the LSU rDNA of T. gondii demonstrated that
XX the D2 domain (C1/C1') region can serve as a target for the
XX development of a species-specific PCR (see AAV72869-75) for the
XX detection of Neospora rDNA, especially in clinical samples. In
XX addition, PCR MIMICS have been developed for use in competitive PCR
XX assays (see AAV72876-81) to differentially diagnose between Neospora
XX spp. and other related/non-related micro-organisms, particularly
XX Toxoplasma spp.
XX
XX Sequence 3499 BP; 870 A; 734 C; 937 G; 958 T; 0 other;
XX
XX Query Match 80.9%; Score 17.8; DB 20; Length 3499;
XX Best Local Similarity 90.5%; Pred. No. 18;
XX Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CCGCTTCAGTCGCCGTACTA 21
XX ||||| ||||| |||||
XX Db 107 CCACTTACTCGCGCTTACTA 87
XX
XX RESULT 14
XX AA088228/C
XX ID AA088228 standard; DNA; 9558 BP.
XX
XX AA088228;
XX
XX 24-NOV-1995 (first entry)
XX
XX Valencia orange ribosomal RNA gene.
XX
XX Ribosomal RNA; rRNA; 18S; 28S; 5.8S; Valencia orange; RFLP analysis;
XX DNA fingerprinting; restriction fragment length polymorphism; ds.
XX
XX Citrus sp.
XX
XX Key
XX FH misc_RNA
XX FT 1..1597
XX FT /tag= a
XX FT /product= 18S_rRNA
XX FT 1598..1848
XX FT /tag= b
XX FT /function= spacer
XX FT 1849..2011
XX FT /tag= c
XX FT /product= 5.8S_rRNA
XX FT 2012..2238
XX FT /tag= d
XX FT /function= spacer
XX FT 2239..5582
XX FT /tag= e
XX

```

Db 222 CGCTCAGTCGCGCTACT 203

RESULT 10

AAV36211/C
ID AAV36211 standard; DNA: 4256 BP.

XX AAV36211;

DT 03-SEP-1998 (first entry)

DE Total contiguous sequence of *P. carinii* in immunosuppressed rats.

KM Detection; diagnosis: 26S rRNA gene; *P. carinii* specific; infection;

KW Species identification; ss.

OS pneumocystis carinii.

PN US5776680-A.

PD 07-JUL-1998.

PP 21-JUL-1995; 95US-0505509.

PR 30-JUL-1992; 92US-0922987.

PR 31-AUG-1994; 94US-0298087.

PA (UYNE-) UNIV NEW JERSEY.

PI Leibowitz MJ, Liu Y;

DR WPI: 1998-398016/34.

PT Detection of *Pneumocystis carinii* - by amplification of nucleic acid
PT from sample with PCR primers specific for the 26S rRNA gene of
PT *Pneumocystis carinii*

XX

PS Disclosure: Columns 29-34: 42pp; English.

XX

CC The present sequence represents the total contiguous sequence of
CC *Pneumocystis carinii* in immunosuppressed Sprague-Dawley rats. The
CC specification describes a method for the diagnosis of *Pneumocystis*
CC *carinii* which comprises detecting the presence of a nucleic acid
CC sequence containing the 26S rRNA gene specific for *P. carinii* in a
CC sample. The 26S rRNA gene in a sample is amplified, and the primer
CC extension products detected by hybridisation with a labelled
CC oligonucleotide. The methods can be used for the diagnosis of
CC *P. carinii* infection and for the detection of various species of
CC *P. carinii*.

SO Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T; 0 other;

Query Match 83.6%; Score 18.4; DR 19; Length 4256;

Best Local Similarity 95.0%; Pred. No. 9.2;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 CGGCTTCACTGCGCTTACT 20

DB 660 CCACTTCACTGCGCTTACT 641

RESULT 11

AAV81001/C

ID AAV81001 standard; DNA: 4256 BP.

XX AAV81001;

DT 23-FEB-1999 (first entry)

DE Contiguous sequence determined for *P. carinii* from immunosuppressed rats.

KW PCR amplification; sequencing; assay: inhibitor; nuclear RNA gene;

KW catalytic Group I self-splicing intron reaction; drug screening; ss.

OS *Pneumocystis carinii*.

XX Key Location/Qualifiers

FT Intron 1..22

FT /tag= a

FT /note= "3'-terminal portion of intron 1"

FT exon 23..53

FT /tag= b

FT /note= "exon 2 of 16S rRNA"

FT misc_feature 54..216

FT /tag= c

FT /note= "internal transcribed spacer 1"

FT exon 217..374

FT /tag= d

FT /note= "5.8S rRNA"

FT misc_feature 375..556

FT /tag= e

FT /note= "internal transcribed spacer 2"

FT exon 557..4256

FT /tag= f

FT /note= "26S rRNA"

PN US5849484-A.

PD 15-DEC-1998.

PP 19-JUN-1995; 95US-0491690.

PR 27-MAY-1993; 93US-0068248.

PR 30-JUL-1992; 92US-0922987.

PR 19-JUN-1995; 95US-0491690.

PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

PI Leibowitz MJ, Liu Y;

DR WPI: 1999-069716/06.

PT Screening assays for drugs against *Pneumocystis carinii* - based on

PT inhibition of 26S rRNA gene intron self-splicing

XX Example 1; fig 2; 51pp; English.

CC The present sequence represents the contiguous sequence determined for
CC *Pneumocystis carinii* from immunosuppressed Sprague-Dawley rats, using
CC the in vitro method of the invention. The method assays for an inhibitor
CC of the catalytic Group I self-splicing intron reaction in the nuclear
CC rRNA genes of *P. carinii*. The method is useful for screening potential
CC drugs for treating *P. carinii* infections before more costly animal
CC testing is conducted.

SO Sequence 4256 BP; 1190 A; 766 C; 1124 G; 1176 T; 0 other;

Query Match 83.6%; Score 18.4; DR 20; Length 4256;

Best Local Similarity 95.0%; Pred. No. 9.2;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 CGGCTTCACTGCGCTTACT 20

DB 660 CCACTTCACTGCGCTTACT 641

RESULT 12

AAV31383

ID AAV31383 standard; DNA: 24 BP.

XX AAV31383;

DT 07-SEP-1998 (first entry)

DE Oligonucleotide Q SEQ ID NO:5 from W09813505.

CC carcinoma, ductal carcinoma in situ or lobular carcinoma in situ. This
 CC sequence represents DNA encoding a human HER polypeptide.
 XX
 SQ Sequence 3040 BP; 851 A; 686 C; 976 G; 527 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 22; Length 3040;
 Best Local Similarity 95.0%; Pred. No. 8.9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCGCTTCACTCGCCGTACT 20
 ||||||||||||
 DB 2780 CCGCTTCACTCGCCGTACT 2799
 RESULT 8
 AA57868
 ID AA57868 standard; cDNA; 3073 BP.
 XX
 AC AA57868;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Protein regulating gene expression PRGE-30 cDNA clone 1977214.
 XX
 KW Protein regulating gene expression; PRGE-30; human;
 KW cell proliferation; antiproliferative; inflammation;
 KW antiinflammatory; gene therapy; diagnosis; RNA binding protein;
 KW RNP-1; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 122..2875
 FT /tag= a
 XX
 PN MO9964596-A2.
 FD 16-DEC-1999.
 XX
 PF 11-JUN-1999; 99WO-US13281.
 XX
 PR 12-JUN-1998; 98US-0089029.
 PR 29-JUL-1998; 98US-0094575.
 PR 14-OCT-1998; 98US-0104624.
 XX
 PA (INCYTE) INCYTE PHARM INC.
 PI Lal P, Yue H, Tang YT, Hillman JL, Bandman O, Corley NC;
 PI Cuegler KJ, Gorgone GA, Baughn MR, Patterson C, Lu DM;
 DR P-PSDB; AAY58637.
 DR WPI: 2000-116543/10.
 XX
 PT New human polypeptides that regulate gene expression, for treatment,
 PT prevention and diagnosis of, e.g. cancer -
 XX
 PS Claim 9; Page 146-147; 150pp; English.
 CC The present sequence is that of Incyte clone 1977214 encoding new
 CC human protein regulating gene expression PRGE-30 (see AAY58637). The
 CC cDNA was initially isolated from pancreatic tumour cDNA library
 CC PACT0102, and the full-length sequence assembled from overlapping
 CC clones from a number of libraries. PRGE-30 is expressed in nervous,
 CC reproductive and gastrointestinal tissues associated with cell
 CC proliferation and inflammation diseases, disorders or conditions.
 CC It is characterised as an RNA binding protein (RNP-1). The invention
 CC provides PRGE polypeptides (see AAY58638) and polynucleotides (see
 CC AA57839-69), expression vectors, host cells, antibodies, agonists and
 CC antagonists. It also provides methods for diagnosing, treating or
 CC preventing disorders associated with expression of PRGE.
 CC Polynucleotides are also used as sources of probes and primers for
 CC diagnosis and monitoring of disease, also for detecting related
 CC sequences and in gene mapping.

XX
 SQ Sequence 3073 BP; 840 A; 699 C; 998 G; 536 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 21; Length 3073;
 Best Local Similarity 95.0%; Pred. No. 9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCGCTTCACTCGCCGTACT 20
 ||||||||||||
 DB 2854 CCGCTTCACTCGCCGTACT 2873
 RESULT 9
 AA192675/C
 ID AA192675 standard; cDNA; 3130 BP.
 XX
 AC AA192675;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 12735.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 DR WPI: 2001-514838/56.
 DR P-PSDB; AA012744.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 1; SEQ ID NO 12735; 1399pp + Sequence listing; English.
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 3130 BP; 566 A; 1005 C; 703 G; 856 T; 0 other;
 Query Match 83.6%; Score 18.4; DB 22; Length 3130;
 Best Local Similarity 95.0%; Pred. No. 9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CCGCTTCACTCGCCGTACT 20

XX DR WPI: 1994-167493/20.
 XX PT Probe for detection of infectious disease - comprises DNA
 PT fragment specific for fungal disease agent, for diagnosis of e.g.
 PT Candida infection
 XX PS Claim 5; Page 19; 35pp; Japanese.
 CC Genomic DNA was isolated from *Candida albicans* (CA-26) (AA063862) and
 CC fragments were generated by an *NotI* digest. This is used as a probe
 CC for detection of *Candida* infection. The advantage of this probe
 CC is that diagnosis time can be reduced from 3-4 days to 1-2, while
 CC maintaining accuracy.
 XX SQ Sequence 899 BP; 233 A; 200 C; 224 G; 242 T; 0 other:
 Query Match 83.6%; Score 18.4; DB 15; Length 899;
 Best Local Similarity 95.0%; Pred. No. 8.1;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 0Y 1 CCGCTTCACTGCGCGTACT 20
 Db 757 CCGCTTCACTGCGCGTACT 738
 RESULT 6
 AAX78193
 ID AAX78193 standard; cDNA; 3040 BP.
 AC AAX78193:
 XX 20-AUG-1999 (first entry)
 DT Human HET cDNA.
 DE
 XX HET; cytosolic; hsp27 promoter; suppressor; hsp27; tumour suppressor;
 KW hsp27-ERE-TAT-binding protein; diagnosis; prognosis; breast cancer;
 KW infiltrating duct carcinoma; lobular carcinoma; medullary carcinoma;
 KW mucinous carcinoma; tubular carcinoma; ductal carcinoma; therapy;
 KW screening; human; ss.
 XX Homo sapiens.
 OS
 XX W09928466-A1.
 PN 10-JUN-1999.
 PD 30-NOV-1998; 98MO-US25381.
 PF 04-DEC-1997; 97US-0068132.
 PR (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX (Allred CD, Fugua SAW, Lee AV, O'Connell P, Oesterreich S;
 PI Osborne CK;
 PI WPI: 1999-385382/32.
 DR P-PSDB: AA008991.
 DR New isolated tumor suppressor, HET
 PT
 XX Claim 3; Page 126-127; 137pp; English.
 XX This invention describes a nucleic acid segment that encodes a novel
 CC human HET polypeptide which has cytosolic activity. The HET protein
 CC binds to the hsp27 promoter in order to suppress expression of the
 CC hsp27 protein. The HET protein (hsp27-ERE-TAT-binding protein) can bind
 CC to the hsp27 promoter and suppress expression of hsp27. The HET acts as a
 CC tumour suppressor gene. The products and methods can be used for the
 CC diagnosis and prognosis of breast cancers including infiltrating duct
 CC carcinoma, lobular carcinoma, medullary carcinoma, mucinous carcinoma,
 CC tubular carcinoma, ductal carcinoma in situ and lobular carcinoma in

CC situ. The products can also be used for therapy and in the screening of
 CC compounds for activity in either stimulating HET activity, overcoming the
 CC lack of HET or blocking the effect of a mutant HET molecule.
 XX SQ Sequence 3040 BP; 851 A; 686 C; 976 G; 527 T; 0 other:
 Query Match 83.6%; Score 18.4; DB 20; Length 3040;
 Best Local Similarity 95.0%; Pred. No. 8.9;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 0Y 1 CCGCTTCACTGCGCGTACT 20
 Db 2780 CCGCTTCACTGCGCGTACT 2799
 RESULT 7
 AAS12099
 ID AAS12099 standard; DNA; 3040 BP.
 XX AAS12099:
 AC AAS12099:
 XX 21-NOV-2001 (first entry)
 DT Human HET DNA sequence.
 DE
 XX HET; nuclear matrix protein; antioestrogen resistance; breast cancer; ds;
 KW scaffold attachment factor B; SAF-B; loss of heterozygosity; human; yuh;
 KW infiltrating duct carcinoma; lobular carcinoma; medullary carcinoma;
 KW mucinous carcinoma; tubular carcinoma; ductal carcinoma;
 KW chromosome 19p13.2-p13.3.
 XX Homo sapiens.
 OS
 XX W0200163292-A2.
 PN 30-AUG-2001.
 PD 22-FEB-2001; 2001MO-US06135.
 PF 22-FEB-2000; 2000US-0184097.
 PR (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Oesterreich S, Osborne CK, Lee AV, Fugua SA;
 PI WPI: 2001-536660/59.
 DR P-PSDB: AA007605.
 DR Detection of antioestrogens breast cancer cells comprises the
 PT measurement of a HET polypeptide -
 PT
 XX Examples: Page 137; 140pp; English.
 XX The invention relates to the detection of resistance to antioestrogens in
 CC breast cancer cells, comprising obtaining a breast cancer cell sample,
 CC contacting the sample with an antibody that specifically binds to the
 CC nuclear matrix protein, HET, (also known as scaffold attachment factor B
 CC or SAF-B) to form a complex and measuring the amount of HET polypeptide
 CC present. A decrease in HET level in the sample compared to the control
 CC indicates a presence of antioestrogen resistance. Resistance in breast
 CC cancer cells can be predicted by comparing the amount of HET protein in
 CC samples from patients with antioestrogen-resistant and
 CC antioestrogen-sensitive breast cancers. Malignant breast cancers can be
 CC diagnosed by determining loss of heterozygosity (LOH) at a chromosome
 CC locus where the LOH at the locus is indicative of antioestrogen
 CC resistance in breast cancer cells, as well as a decrease in chance of
 CC survival. The cancerous cells may be from an infiltrating duct carcinoma,
 CC lobular carcinoma, medullary carcinoma, mucinous carcinoma, tubular

PR 24-NOV-1987; 87WO-US03009.
 PR 11-DEC-1991; 91US-0806929.
 PR 24-NOV-1986; 86US-0934244.
 PR 07-AUG-1987; 87US-0083542.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI McDonough SH, Kop JA, Smth RD, Hogan JJ;
 DR WPI: 2001-060029/07.
 XX
 PT Preparing a probe for nucleic acid hybridization assays comprises
 PT constructing a nucleotide polymer sufficiently complementary to
 PT hybridize to an rRNA region that distinguishes non-viral target from
 PT non-viral non-target species -
 XX
 PS Disclosure; Fig 5; 75pp; English.
 CC
 CC The present invention provides novel methods of producing probes for use
 CC in the identification of a number of microorganisms. These include E.
 CC coli, Mycobacteria, Mycoplasma, Campylobacter, Chlamydia, Enterobacter,
 CC Legionella, Salmonella, Pseudomonas, Neisseria gonorrhoeae, fungi and
 CC bacteria.
 CC
 SO Sequence 3550 BP; 941 A; 697 C; 1000 G; 912 U; 0 other;
 Query Match 95.5%; Score 21; DB 22; Length 3550;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGCTTCACTGCGCGTTACTA 21 22
 DB 261 CCGCTTCACTGCGCGTTACTA 241
 RESULT 4
 AAQ39050/C
 ID AAQ39050 standard; DNA; 6824 BP.
 XX
 AC AAQ39050;
 XX
 DT 28-JUL-1993 (first entry)
 XX
 DE K.lactis/S. cerevisiae genetic vector.
 XX
 KM Genetic; vector; integration; Kluyveromyces lactis; 255 ribosomal DNA;
 KM Saccharomyces cerevisiae; E. coli; domain; yeast; plasmid; promoter;
 KM expression cassette; HIS3; marker; transformant; human; lysozyme; HLZ;
 KM GAL7; signal sequence; killer toxin; transcription termination signal;
 KM FLP; 2 micron plasmid; ss.
 XX
 OS Synthetic.
 XX
 PN EP537456-A.
 XX
 PD 21-APR-1993.
 XX
 PF 31-AUG-1992; 92EP-0114838.
 XX
 PR 04-SEP-1991; 91IT-0M12349.
 XX
 PA (ISTS) SCLAVO SPA.
 XX
 PI Galeotti CL, Gallo E, Riccio ML, Rossolini GM, Thaller MC;
 DR WPI: 1993-127394/16.
 XX
 PT Vector for Kluyveromyces lactis and Saccharomyces cerevisiae -
 PT which allows stable multiple integration of DNA for prodn. of
 PT heterologous proteins
 XX
 PS Claim 1; Fig 1; 26pp; English.

CC This sequence represents a genetic vector which allows the stable
 CC multiple integration of DNA sequences into the genome of Kluyveromyces
 CC lactis and Saccharomyces cerevisiae. This sequence can be used in an
 CC integrating vector which comprises a region necessary for the stable
 CC maintenance of the plasmid in E. coli and a domain which acts as an
 CC integrating unit consisting of two not contiguous sequences of the 255
 CC ribosomal DNA from S. cerevisiae, flanking a genetic marker suitable
 CC for selection of the yeast transformants in which the integration
 CC event has occurred. Other DNA sequences may be introduced into the
 CC integration plasmid, such as expression cassettes. The gene HIS3
 CC from K. lactis and S. cerevisiae is pre. used as a genetic marker
 CC for the selection of transformants and an expression cassette for the
 CC production and secretion into the culture medium of human lysozyme.
 CC This complete transformation vector is 7850 bp long and includes the
 CC integration vector of the invention and an expression cassette
 CC comprising the K. lactis GAL7 promoter, the signal sequence of the K.
 CC lactis killer toxin, the cDNA encoding the ripe form of human lysozyme
 CC (HLZ) and the transcription termination signal FLP of the 2 micron
 CC plasmid from S. cerevisiae.
 CC
 SO Sequence 6824 BP; 1815 A; 1521 C; 1726 G; 1762 T; 0 other;
 Query Match 95.5%; Score 21; DB 14; Length 6824;
 Best Local Similarity 100.0%; Pred. No. 0.48;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGCTTCACTGCGCGTTACTA 21
 DB 144 CCGCTTCACTGCGCGTTACTA 124
 RESULT 5
 AAQ63962/C
 ID AAQ63962 standard; DNA; 699 BP.
 XX
 AC AAQ63962;
 XX
 DT 30-NOV-1994 (first entry)
 XX
 DE Candida albicans (CA-26) genomic DNA fragment.
 XX
 KM Detection of infectious disease; fungal disease agent;
 KM probes; diagnosis; ds.
 XX
 OS Candida albicans.
 XX
 FH Key Location/Qualifiers
 FT misc-feature 14..202
 FT /tag- a
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 FT misc-feature 372..737
 FT /tag- b
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 FT misc-feature 517..737
 FT /tag- c
 FT /note- "fragment isolated and used as a template
 FT for probe formation"
 PN WO9410341-A.
 XX
 PD 11-MAY-1994.
 XX
 PF 25-OCT-1993; 93WO-JP01555.
 XX
 PR 23-OCT-1992; 92JP-0285802.
 XX
 PA (FUSO) FUSO YAKUHIN KOGYO KK.
 PA (ONOV/) ONO Y.
 PA (FUSO) FUSO PHARM IND LTD.
 PI Hirotsu T, Keshi H, Matsuhisa A, Ohno T;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 07:01:56 ; Search time 141.778 Seconds

(without alignments)
266.418 Million cell updates/sec

Title: US-09-780-113D-8

Perfect score: 22

Sequence: 1 ccgccttaactccgccttaactag 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Capext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	22	100.0	218	20	AAZ33745
C 2	22	100.0	431	21	AAF09568
C 3	21	95.5	3550	22	AAF23019
C 4	21	95.5	6824	14	AAQ39050
C 5	18.4	83.6	899	15	AAQ63962
C 6	18.4	83.6	3040	20	AAZ78193
C 7	18.4	83.6	3040	22	AAZ12099
C 8	18.4	83.6	3073	21	AAZ57868
C 9	18.4	83.6	3130	22	AAI92675

C 10	18.4	83.6	4256	19	AAV36211	Total contiguous s
C 11	18.4	83.6	4256	20	AAV83003	Contiguous sequenc
C 12	17.8	80.9	24	19	AAV31383	Oligonucleotide Q
C 13	17.8	80.9	3499	20	AAV72868	Neospora caninum 1
C 14	17.8	80.9	9558	16	AAQ88228	Valencia orange ri
C 15	17.8	80.9	59590	21	AAZ22281	BAC containing rep
C 16	16.8	76.4	23	22	AAZ76254	Cordyceps sinensis
C 17	16.8	76.4	415	23	AAZ71954	DNA encoding novel
C 18	16.8	76.4	636	21	AAZ22024	Human breast and o
C 19	16.8	76.4	1482	13	AAQ25896	ACC synthase gen
C 20	16.8	76.4	1703	12	AAQ15131	Clone pACT1 encodi
C 21	16.8	76.4	1703	19	AAV15701	Zucchini ACC synth
C 22	16.8	76.4	1703	22	AAAD04541	Zucchini 1-aminocy
C 23	16.8	76.4	1703	22	AAZ3618	Zucchini ACC synth
C 24	16.8	76.4	9060	12	AAQ15132	Zucchini ACC synth
C 25	16.8	76.4	9060	19	AAV15702	Zucchini ACC synth
C 26	16.8	76.4	9060	22	AAAD04542	Zucchini CP-ACC 1A
C 27	16.8	76.4	9060	22	AAZ23619	Zucchini ACC synth
C 28	16.8	76.4	13076	22	AAZ06214	Human reproductive
C 29	16.8	76.4	13076	22	AAZ07355	Human reproductive
C 30	16.2	73.6	233	18	AAV78892	Staphylococcus aur
C 31	16.2	73.6	265	18	AAV78804	Staphylococcus aur
C 32	16.2	73.6	313	22	AAH50814	Human tumour assoc
C 33	16.2	73.6	324	22	AAH50846	Human tumour assoc
C 34	16.2	73.6	370	18	AAV78337	Staphylococcus aur
C 35	16.2	73.6	390	18	AAV78363	Staphylococcus aur
C 36	16.2	73.6	400	18	AAV78215	Staphylococcus aur
C 37	16.2	73.6	400	18	AAV78098	Staphylococcus aur
C 38	16.2	73.6	400	18	AAV77867	Staphylococcus aur
C 39	16.2	73.6	518	18	AAV77864	Staphylococcus aur
C 40	16.2	73.6	1044	18	AAV74959	Staphylococcus aur
C 41	16.2	73.6	1118	18	AAV77847	Staphylococcus aur
C 42	16.2	73.6	2923	22	AAZ89402	S.aureus 23S rRNA
C 43	16.2	73.6	5715	23	ABL12428	Drosophila melanog
C 44	16.2	73.6	5715	23	ABL17526	Drosophila melanog
C 45	16.2	73.6	6591	18	AAV77425	Staphylococcus aur

ALIGNMENTS

RESULT 1	AAZ33745/c
ID	AAZ33745 standard; cDNA, 218 BP.
XX	AAZ33745;
AC	
XX	
DT	09-DEC-1999 (first entry)
XX	
DE	Tobacco plant resistance-associated cDNA fragment 70.
XX	
KW	Tobacco; plant; resistance; tobacco mosaic virus; TMV; infection;
KW	protection; plant protection agent; phytopathogenic fungi; nematode;
KW	insect pest; pathogen resistance; transgenic plant; diagnosis; ss.
XX	
OS	Nicotiana tabacum.
XX	
PN	DE19813048-A1.
XX	
PD	30-SEP-1999.
XX	
PF	25-MAR-1998; 98DE-1013048.
XX	
PR	25-MAR-1998; 98DE-1013048.
XX	
PA	(FARB) BAYER AG.
XX	
PI	Schneider P, Garbers C, Langen G, Kiedrowski S;
XX	
DR	WPI: 1999-552163/47.
XX	
PT	Nucleic acids useful for identifying compounds capable of inducing
PT	pathogen resistance in plants

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 07:07:41 ; Search time 31.2889 Seconds

(without alignments)
172.711 Million cell updates/sec

Title: US-09-780-113d-8

Perfect score: 22

Sequence: 1 ccgcctcactccgcctcactag 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	95.5	666	4	US-08-998-416-1013
2	21	95.5	697	4	US-08-998-416-909
3	21	95.5	699	4	US-08-998-416-774
4	21	95.5	700	4	US-08-998-416-813
5	21	95.5	720	4	US-08-998-416-601
6	18.4	83.6	899	1	US-08-416-831B-1
7	18.4	83.6	4256	2	US-08-505-509-31
8	18.4	76.4	1703	4	US-08-491-690A-31
9	16.8	76.4	9060	4	US-08-378-313-18
10	16.8	76.4	9060	4	US-08-378-313-20
11	16	72.7	33	1	US-08-308-461-3
12	16	72.7	33	5	PCT-US95-11823-3
13	15.4	70.0	1627	3	US-09-106-464-1
14	15	68.2	20	2	US-08-778-912A-9
15	15	68.2	20	2	US-08-778-912A-11
16	15	68.2	20	2	US-08-778-912A-15
17	15	68.2	20	4	US-09-541-941B-9
18	15	68.2	20	4	US-09-541-941B-11
19	15	68.2	20	4	US-09-541-941B-15
20	15	68.2	815	4	US-09-541-941B-20
21	15	68.2	870	2	US-08-778-912A-4
22	15	68.2	874	2	US-09-541-941B-4
23	15	68.2	874	2	US-08-778-912A-1
24	15	68.2	874	4	US-09-541-941B-2
25	15	68.2	875	2	US-08-778-912A-2
26	15	68.2	875	2	US-08-778-912A-3
27	15	68.2	875	2	US-08-778-912A-5

28	15	68.2	875	2	US-08-778-912A-7	Sequence 7, Appl 1
29	15	68.2	875	4	US-09-541-941B-1	Sequence 1, Appl 1
30	15	68.2	875	4	US-09-541-941B-3	Sequence 3, Appl 1
31	15	68.2	875	4	US-09-541-941B-5	Sequence 5, Appl 1
32	15	68.2	875	4	US-09-541-941B-7	Sequence 7, Appl 1
33	15	68.2	876	2	US-08-778-912A-6	Sequence 6, Appl 1
34	15	68.2	876	4	US-09-541-941B-6	Sequence 6, Appl 1
35	15	68.2	898	4	US-09-541-941B-22	Sequence 22, Appl 1
36	15	68.2	898	4	US-09-541-941B-23	Sequence 23, Appl 1
37	15	68.2	898	4	US-09-541-941B-24	Sequence 24, Appl 1
38	15	68.2	898	4	US-09-541-941B-25	Sequence 25, Appl 1
39	15	68.2	904	4	US-09-541-941B-21	Sequence 21, Appl 1
40	15	68.2	917	4	US-09-541-941B-18	Sequence 18, Appl 1
41	15	68.2	918	4	US-09-541-941B-16	Sequence 16, Appl 1
42	15	68.2	918	4	US-09-541-941B-17	Sequence 17, Appl 1
43	15	68.2	918	4	US-09-541-941B-19	Sequence 19, Appl 1
44	14.6	66.4	681	1	US-08-378-588-20	Sequence 20, Appl 1
45	14.6	66.4	681	2	US-08-811-094-20	Sequence 20, Appl 1

ALIGNMENTS

RESULT 1
US-08-998-416-1013/c
Sequence 1013, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Red: schung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC 27709
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8689
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1013:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PA61621UP

US-08-998-416-1013

Query Match 95.5%; Score 21; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 457 CCGCTTCACTCGCCGTACTA 437

RESULT 2

US-08-998-416-909
; Sequence 909, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 909:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1567RP
; US-08-998-416-909

Query Match 95.5%; Score 21; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 332 CCGCTTCACTCGCCGTACTA 352

RESULT 3

US-08-998-416-774/c

; Sequence 774, Application US/08998416
; Patent No. 6239264

; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 774:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1498BP
; US-08-998-416-774

Query Match 95.5%; Score 21; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCCGTACTA 21
|||||
DB 457 CCGCTTCACTCGCCGTACTA 437

RESULT 4

US-08-998-416-813
; Sequence 813, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Redischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII

TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 813:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1518RP
US-08-998-416-813

Query Match 95.5%; Score 21; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGTTCACTCGCGTTACTA 21
|||||
Db 332 CGCGTTCACTCGCGTTACTA 352

RESULT 5
US-08-998-416-601
Sequence 601 Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 601:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1411RP
US-08-998-416-601

Query Match 95.5%; Score 21; DB 4; Length 720;
Best Local Similarity 100.0%; Pred. No. 0.058;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCGTTCACTCGCGTTACTA 21
|||||
Db 332 CGCGTTCACTCGCGTTACTA 352

RESULT 6
US-08-416-831B-1/c
Sequence 1, Application US/08416831B
Patent No. 5708159
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Hirotsu, Takuo
APPLICANT: Keshi, Hiroyuki
APPLICANT: Matsuhisa, Akio
TITLE OF INVENTION: Probe for Diagnosing Infectious Diseases
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,831B
FILING DATE: 19-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/01555
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPA 4-285802
FILING DATE: 23-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 19036/32578
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 899 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Candida albicans
STRAIN: Clinical isolate CA-26
US-08-416-831B-1

Query Match 83.6%; Score 18.4; DB 1; Length 899;
Best Local Similarity 95.0%; Pred. No. 1.2;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 757 CCGCTTCACTCCCGCTTACT 738

RESULT 7
US-08-505-509-31/C
Sequence 31, Application US/08505509
Patent No. 577680
GENERAL INFORMATION:
APPLICANT: Liebowitz, Michael J.
APPLICANT: Liu, Yong
TITLE OF INVENTION: Diagnostic Probes for
TITLE OF INVENTION: Pneumocystis Carinii
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: P.O. Box 1267
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08551
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,509
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/298,087
FILING DATE:
APPLICATION NUMBER: US/07/922,987
FILING DATE: 30-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD1-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 466-3407
TELEFAX: (609) 466-2760
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4256 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-505-509-31

Query Match 83.6%; Score 18.4; DB 1; Length 4256;
Best Local Similarity 95.0%; Pred. No. 1.4;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 660 CCGCTTCACTCCCGCTTACT 641

RESULT 8
US-08-491-690A-31/C
Sequence 31, Application US/08491690A
Patent No. 5849484
GENERAL INFORMATION:
APPLICANT: Liebowitz, Michael J.
APPLICANT: Liu, Yong
TITLE OF INVENTION: In Vitro Assay For Inhibitors
TITLE OF INVENTION: Of The Intron Self-Splicing Reaction in Pneumocystis Carinii
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: 758 Springfield Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,690A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,248
FILING DATE: 27-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD1-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 273-4988
TELEFAX: (908) 273-4679
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 4256 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-491-690A-31

Query Match 83.6%; Score 18.4; DB 2; Length 4256;
Best Local Similarity 95.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCCCGCTTACT 20
|||||
Db 660 CCGCTTCACTCCCGCTTACT 641

RESULT 9
US-08-378-313-18
Sequence 18, Application US/08378313
Patent No. 6207881
GENERAL INFORMATION:
APPLICANT: THEOLOGIS, ATHANASIOS
APPLICANT: SATO, TAKAHITO
TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH
TITLE OF INVENTION: GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,313
FILING DATE:
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/862,493
FILING DATE: 02-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 29190-20002.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 856-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1703 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 11..1489
US-08-378-313-18

Query Match 76.4%; Score 16.8; DB 4; Length 1703;
Best Local Similarity 90.0%; Pred. No. 8.1;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CGCTTCACTCGCCGTACTA 21
DB 1458 CGCCTCCTACTCGCCGTACTA 1477

RESULT 10
US-08-378-313-20
Sequence 20, Application US/08378313
Patent No. 6207881
GENERAL INFORMATION:
APPLICANT: THEOLOGIS, ATHANASIOS
APPLICANT: SATO, TAKAHIDO
TITLE OF INVENTION: CONTROL OF FRUIT RIPENING THROUGH
NUMBER OF SEQUENCES: 34
TITLE OF INVENTION: GENETIC CONTROL OF ACC SYNTHASE SYNTHESIS
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,313
FILING DATE:
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/862,493

FILING DATE: 02-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 29190-20002.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 856-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 9060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: join(2704..2880, 2968..3099, 3183..3344, 3810
LOCATION: ..4376, 4463..4903)
US-08-378-313-20

Query Match 76.4%; Score 16.8; DB 4; Length 9060;
Best Local Similarity 90.0%; Pred. No. 9.7;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CGCTTCACTCGCCGTACTA 21
DB 4872 CGCCTCCTACTCGCCGTACTA 4891

RESULT 11
US-08-308-461-3
Sequence 3, Application US/08308461
Patent No. 5706640
GENERAL INFORMATION:
APPLICANT: Jeffrey W. Voss
APPLICANT: Connie Caron
TITLE OF INVENTION: Inducers of Gamma Globin Gene Expression
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/308,461
FILING DATE: 16-SEP-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: B91-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: oligonucleotide
US-08-308-461-3

Query Match 72.7%; Score 16; DB 1; Length 33;

Best Local Similarity 100.0%; Pred. No. 14;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TTCACGCGCGTACT 20
|||||
Db 1 TTCACGCGCGTACT 16

RESULT 12

PCT-US95-11823-3

Sequence 3, Application PC/US9511823

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Inducers of Gamma Globin Gene Expression

TITLE OF INVENTION: and Screening Assays Therefor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahive & Cockfield

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/11823

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/308,461

FILING DATE: 16-SEP-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: DeConti, Giulio A. Jr.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: BBI-015PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: oligonucleotide

PCT-US95-11823-3

Query Match 72.7%; Score 16; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TTCACGCGCGTACT 20
|||||
Db 1 TTCACGCGCGTACT 16

RESULT 13

US-09-106-464-1/c

Sequence 1, Application US/09106464

Patent No. 601145

GENERAL INFORMATION:

APPLICANT: Steffens, John C.

APPLICANT: Shengas, Gurdev S.

APPLICANT: Koval, Jian-Ping

APPLICANT: Bannetta, Nancy

TITLE OF INVENTION: Chain length specific UDP-Glc:Fatly Acid

TITLE OF INVENTION: Glucosyltransferases

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones, Tullar & Cooper, P.C.

STREET: P.O. Box 2266 Eads Station

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/106,464

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/055,554

FILING DATE: 13-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: Spector, Eric S.

REGISTRATION NUMBER: 22495

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-415-1500

TELEFAX: 703-415-1508

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1627 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1413

US-09-106-464-1

Query Match 70.0%; Score 15.4; DB 3; Length 1627;
Best Local Similarity 94.1%; Pred. No. 41;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCACGCGCGTACT 20
|||||
Db 1291 CTCACGCGCGTACT 1275

RESULT 14

US-08-778-912A-9

Sequence 9, Application US/08778912A

Patent No. 5876977

GENERAL INFORMATION:

APPLICANT: Wang, Jun

APPLICANT: Ngan, F N

APPLICANT: But, P P

APPLICANT: Shaw, P C

TITLE OF INVENTION: POLYMERASE CHAIN REACTION-RESTRICTION FRAGMENT

TITLE OF INVENTION: LENGTH POLYMORPHISM TEST FOR THE AUTHENTICATION OF

TITLE OF INVENTION: TRADITIONAL CHINESE MEDICINES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/778,912A

FILED DATE: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chan, Albert Wai-Kit
REGISTRATION NUMBER: 36,479
REFERENCE/DOCKET NUMBER: 52188
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400
TELEFAX: 212 391 0526
INFORMATION FOR SEQ. ID NO.: 9
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-778-912A-9

Query Match 68.2%; Score 15; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 ACTGCCGCTTACTAG 22
DB 1 ACTGCCGCTTACTAG 15

RESULT 15
US-08-778-912A-11
Sequence 11, Application US/0878912A
Patent No. 5876977
GENERAL INFORMATION:
APPLICANT: Wang, Jun
APPLICANT: Ngan, F N
APPLICANT: But, P P
APPLICANT: Shaw, P C
TITLE OF INVENTION: POLYMERASE CHAIN REACTION-RESTRICTION FRAGMENT
TITLE OF INVENTION: LENGTH POLYMORPHISM TEST FOR THE AUTHENTICATION OF
TITLE OF INVENTION: TRADITIONAL CHINESE MEDICINES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,912A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chan, Albert Wai-Kit
REGISTRATION NUMBER: 36,479
REFERENCE/DOCKET NUMBER: 52188
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 278 0400
TELEFAX: 212 391 0526
INFORMATION FOR SEQ. ID NO.: 11
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-778-912A-11

Query Match

68.2%; Score 15; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ACTGCCGCTTACTAG 22
DB 1 ACTGCCGCTTACTAG 15

Search completed: October 17, 2002, 08:51:48
Job time: 33.2889 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:24:35 ; Search time 1156.22 Seconds
(without alignments)
256.813 Million cell updates/sec

Title: US-09-780-113D-8

Perfect score: 22

Sequence: 1 ccgcctcactgcgcctactag 22

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_estlin:*
9: em_estlin:*
10: em_estlin:*
11: em_estlin:*
12: em_estlin:*
13: em_estlin:*
14: em_estlin:*
15: em_estlin:*
16: em_estlin:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	309	9	A1904541 PM-BT057-
2	22	100.0	311	9	AW724745 f7h10nm.f
3	22	100.0	312	9	A1320035 B8902nm.f
4	22	100.0	313	10	BE776149 MY-11-E-0
5	22	100.0	314	9	A1904502 PM-BT057-
6	22	100.0	359	9	AW981798 PC18H03 P
7	22	100.0	359	9	AW982034 PC03D12 P
8	22	100.0	367	9	AW754545 PC03D12 P
9	22	100.0	517	10	BE356832 DG1_145_D
10	21	95.5	206	12	A2925177 4910_e232
11	21	95.5	232	12	A2917057 4911_fd63
12	21	95.5	262	12	A2932421 474_dh298
13	21	95.5	273	12	A2932397 474_dh298
14	21	95.5	319	12	A2916987 4911_fd63
15	21	95.5	330	12	A2124480 K025r S.
16	21	95.5	369	12	A2925656 4910_e234
17	21	95.5	400	12	A2924119 4906_1c27

18	21	95.5	427	10	T39104 EST1046/5 S
19	21	95.5	432	12	A2923503 4908_4f21
20	21	95.5	437	12	A2917126 4911_4d64
21	21	95.5	454	12	CNS07850
22	21	95.5	459	12	AQ501835 V13H11 MT
23	21	95.5	469	12	AQ0875401 V124H8 MT
24	21	95.5	482	12	AQ0875198 V12366 MT
25	21	95.5	488	12	A2923389 4908_4f20
26	21	95.5	493	12	AQ0874532 V110B9 MT
27	21	95.5	506	12	A2930728 474_dh57
28	21	95.5	514	12	A2931326 474_dh64
29	21	95.5	516	12	A2928557 479_dh111
30	21	95.5	516	12	A2928946 479_dh120
31	21	95.5	517	12	A2928988 479_dh126
32	21	95.5	518	12	A2929426 479_dh108
33	21	95.5	519	12	A2928336 479_dh108
34	21	95.5	519	12	A2928515 479_dh110
35	21	95.5	519	12	A2929334 479_dh125
36	21	95.5	520	12	A2930156 474_dh50
37	21	95.5	520	12	A2928166 479_dh105
38	21	95.5	520	12	A2928211 479_dh105
39	21	95.5	520	12	A2929053 479_dh121
40	21	95.5	525	12	AQ0875993 V13H19 MT
41	21	95.5	532	12	AL394655 13_end OF
42	21	95.5	542	12	A2929712 479_dh188
43	21	95.5	542	12	AQ502375 V2704 mtn
44	21	95.5	654	12	AQ492180 V104H11 m
45	21	95.5	664	12	AL436874 13_end OF

ALIGNMENTS

RESUT.F 1
A1904541/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 309)
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
Nagai M.A., de Silva W. Jr., Zago M.A., Bordin S., Costa I.F.,
Goldman G.H., Carvalho A.F., Matsukuma A., Bala G.S., Simpson L.H.,
Brunstein A., de Oliveira P.S., Bucher P., Jongeneel C.V., Chare
, M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
Simpson A.J.

REFERENCE

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE

JOURNAL MEDLINE COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?cl=PM62-PM-BT057-305.html
Seq primer: puc 18 forward.
Location/Qualifiers
1..309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT057"

FEATURES SOURCE

/sex="female"
/dev.stage="Adult"
/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 87 a 67 c 90 g 65 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 88 CCGCTTCACTGCGCGTTACTAG 67

RESULT 2
LOCUS AM724745 311 bp mRNA linear EST 19-APR-2000
DEFINITION f7h10nm.f1 Neurospora crassa morning cDNA library Neurospora crassa
ACCESSION AM724745
VERSION AM724745.1 GI:7619305
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 311)
AUTHORS Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
TITLE Two Neurospora crassa EST Databases
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: Universal Forward Primer
High quality sequence stop: 266.

FEATURES
Location/Qualifiers
1..311
source

/organism="Neurospora crassa"
/strain="bd, frg7 A"
/db_xref="taxon:5141"
/clone="f7h10nm"
/clone_1b="Neurospora crassa morning cDNA library"
/tissue_type="tissue harvested following 22hr growth in dark"
/note="Vector: pBluescript SK-; Site.1: Xba1; Site.2: EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.
5' end of cDNA cloned into EcoRI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 61 a 100 c 69 g 81 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 158 CCGCTTCACTGCGCGTTACTAG 179

RESULT 3
LOCUS AI320035 312 bp mRNA linear EST 18-DEC-1998
DEFINITION b8902nm.f1 Neurospora crassa morning cDNA library Neurospora crassa
ACCESSION AI320035
VERSION AI320035.1 GI:4036017
KEYWORDS EST.
SOURCE Neurospora crassa.
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 312)
AUTHORS Zhu,H., Lai,H., Kupfer,D., Dunlap,J.C. and Roe,B.A.
TITLE Two Neurospora crassa EST Databases
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: Universal Forward Primer
High quality sequence stop: 262.

FEATURES
Location/Qualifiers
1..312
source

/organism="Neurospora crassa"
/strain="bd, frg7 A"
/db_xref="taxon:5141"
/clone="b8902nm"
/clone_1b="Neurospora crassa morning cDNA library"
/tissue_type="tissue harvested following 22hr growth in dark"
/note="Vector: pBluescript SK-; Site.1: Xba1; Site.2: EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.
5' end of cDNA cloned into Xba1 site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 63 a 101 c 67 g 81 t
ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 158 CCGCTTCACTGCGCGTTACTAG 179

RESULT 4
LOCUS BE776149/c 313 bp mRNA linear EST 20-SEP-2000
DEFINITION MY-11-E-03 Pinifastansmy Phytophthora infestans cDNA, mRNA sequence.
ACCESSION BE776149
VERSION BE776149.1 GI:10229804
KEYWORDS EST.
SOURCE potato late blight agent.
ORGANISM Phytophthora infestans
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
Phytophthora.

REFERENCE 1 (bases 1 to 313)
AUTHORS Kamun,S., Haeber,P.T., Sobral,B.W.S., Nuss,D. and Govers,F.
TITLE Initial assessment of gene diversity for the oomycete pathogen
Phytophthora infestans based on expressed sequences
JOURNAL Fungal Genet. Biol. 28 (2), 94-106 (1999)
MEDLINE 20056376
COMMENT Contact: Govers F
Laboratory of Phytopathology
Wageningen University
Bilthoven 9, P.O.Box 8025, 6700 EE, Wageningen, The Netherlands

us-09-780-113d-8.rst

545

```

/db_xref=
/clone_lib="BT057"
sex="female"
sex="male"
site_1: sma1; site_2:

```

QY	Db
1	103
CCGCTTCACTGCGGCTTACTAG	CCGCTTCACTGCGGCTTACTAG

Mon Oct 21 16:44:31 2002

us-09-780-113d-8.1st

LACTUS
DEFINITION AM982034
ACCESSION PC23A10 Pine Triplex pollen cone 359 bp mRNA
VERSION AM982034 mRNA sequence.
KEYWORDS AM982034.1
SOURCE EST
ORGANISM 10b1011y pine.
Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus taeda

REFERENCE
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
JOURNAL The Pine Gene Discovery Project
COMMENT Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology
North Carolina State University
Dept. of Forestry, NC State University
Tel: 27695-8008
Fax: 919-515-7800
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 359
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC23A10"
/dev_stage="Pine Triplex pollen cone library"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1; Isolation: in the early spring, frozen and used for mRNA preparation. The SMART-PCR method (Clontech) was used to recover a library from 1 ug total RNA, using the lambda Triplex vector. Plasmid subclones in pTriplex were sequenced from the 5' end." 83 t

BASE COUNT
ORIGIN 101 a
73 c 102 g 83 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 359;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCGCTTCACTCGCGCTTACTAG 22
|||||
103 CCGCTTCACTCGCGCTTACTAG 22

ULT 8
54545/c
US
INITIATION AM754545
PC03D12 Pine Triplex pollen cone library Pinus taeda cDNA clone
SSION PC03D12, mRNA sequence.
ION AM754545 367 bp
RDS AM754545.1 GI:7676265
E. EST. EST 01-MAY-2000
ANISM 10b1011y pine.
Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus taeda
REFERENCE
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
JOURNAL The Pine Gene Discovery Project
COMMENT Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology
North Carolina State University
Dept. of Forestry, NC State University
Tel: 27695-8008
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 359
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC03D12"
/dev_stage="Pine Triplex pollen cone library"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1; Isolation: in the early spring, frozen and used for mRNA preparation. The SMART-PCR method (Clontech) was used to recover a library from 1 ug total RNA, using the lambda Triplex vector. Plasmid subclones in pTriplex were sequenced from the 5' end." 83 t

Page 4

FEATURES
source
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 367
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC03D12"
/dev_stage="Pine Triplex pollen cone library"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1; Isolation: in the early spring, frozen and used for mRNA preparation. The SMART-PCR method (Clontech) was used to recover a library from 1 ug total RNA, using the lambda Triplex vector. Plasmid subclones in pTriplex were sequenced from the 5' end." 84 t

BASE COUNT
ORIGIN 99 a
70 c 105 g 84 t
Query Match
Best Local Similarity 100.0%; Score 22; DB 9; Length 367;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCGCTTCACTCGCGCTTACTAG 22
|||||
103 CCGCTTCACTCGCGCTTACTAG 22

RESULT 9
BE356832/c
LOCUS BE356832.1
DEFINITION DC1_145.D09.B1.A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
ACCESSION BE356832
VERSION BE356832
KEYWORDS EST
SOURCE EST: 9298389
ORGANISM Sorghum.
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PAC
REFERENCE
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
JOURNAL The Pine Gene Discovery Project
COMMENT Unpublished (1999)
Contact: Ross Whetten
Forest Biotechnology
North Carolina State University
Dept. of Forestry, NC State University
Tel: 27695-8008
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 367
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC03D12"
/dev_stage="Pine Triplex pollen cone library"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1; Isolation: in the early spring, frozen and used for mRNA preparation. The SMART-PCR method (Clontech) was used to recover a library from 1 ug total RNA, using the lambda Triplex vector. Plasmid subclones in pTriplex were sequenced from the 5' end." 84 t

FEATURES
source
Email: rosswhetten@ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 367
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone.lib="PC03D12"
/dev_stage="Pine Triplex pollen cone library"
/lab_host="E. coli BM25.8"
/note="Organ: pollen cone; Vector: Lambda Triplex; Site: 1; Isolation: in the early spring, frozen and used for mRNA preparation. The SMART-PCR method (Clontech) was used to recover a library from 1 ug total RNA, using the lambda Triplex vector. Plasmid subclones in pTriplex were sequenced from the 5' end." 84 t

Query Match 100.0%; Score 22; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCCGTTACTAG 22
|||||
Db 75 CCGCTTCACTCGCCGTTACTAG 54

RESULT 10
AZ925177 206 bp DNA linear GSS 01-APR-2001
LOCUS
DEFINITION 4910.ez32908.s1 Saccharomyces paradoxus N17 Saccharomyces paradoxus genomic clone 4910.ez32908.s1, DNA sequence.
ACCESSION AZ925177
VERSION AZ925177.1 GI:13496076
KEYWORDS GSS.
SOURCE Saccharomyces paradoxus.
ORGANISM Saccharomyces paradoxus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 206)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis
JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mjgenetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..206
/organism="Saccharomyces paradoxus"
/strain="N17"
/db_xref="taxon:27291"
/clone="4910.ez32908.s1"
/clone_lib="Saccharomyces paradoxus N17"
/note="Random genomic sequence"

BASE COUNT 51 a 64 c 39 g 52 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 206;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCCGTTACTA 21
|||||
Db 164 CCGCTTCACTCGCCGTTACTA 184

RESULT 11
AZ917057 232 bp DNA linear GSS 01-APR-2001
LOCUS
DEFINITION 4911.f63n01.s1 Saccharomyces bayanus MCYC 623-6C Saccharomyces bayanus genomic clone 4911.f63n01.s1, DNA sequence.
ACCESSION AZ917057
VERSION AZ917057.1 GI:13493581
KEYWORDS GSS.
SOURCE Saccharomyces bayanus.
ORGANISM Saccharomyces bayanus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 232)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis

JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mjgenetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..232
/organism="Saccharomyces bayanus"
/strain="MCYC 623-6C"
/db_xref="taxon:4931"
/clone="4911.f63n01.s1"
/clone_lib="Saccharomyces bayanus MCYC 623-6C"
/note="Random genomic sequence"

BASE COUNT 68 a 44 c 62 g 58 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 232;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCCGTTACTA 21
|||||
Db 112 CCGCTTCACTCGCCGTTACTA 92

RESULT 12
AZ932421 262 bp DNA linear GSS 01-APR-2001
LOCUS
DEFINITION 474.dhz9g11.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces unisporus genomic clone 474.dhz9g11.s1, DNA sequence.
ACCESSION AZ932421
VERSION AZ932421.1 GI:13503334
KEYWORDS GSS.
SOURCE Saccharomyces unisporus.
ORGANISM Saccharomyces unisporus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 262)
AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.
TITLE Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis
JOURNAL Unpublished (2001)
COMMENT Contact: Johnston M
Department of Genetics
Washington University Medical School
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
Tel: 314 362 2735
Fax: 314 362 7855
Email: mjgenetics.wustl.edu
Class: random plasmid subclone.
FEATURES
source
1..262
/organism="Saccharomyces unisporus"
/strain="NRRL Y-1556 (CBS 398)"
/db_xref="taxon:27294"
/clone="474.dhz9g11.s1"
/clone_lib="Saccharomyces unisporus NRRL Y-1556"
/note="Random genomic sequence"

BASE COUNT 69 a 84 c 43 g 66 t
ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 262;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCCGTTACTA 21
|||||
Db 211 CCGCTTCACTCGCCGTTACTA 231

RESULT 13

LOCUS A2932397 273 bp DNA linear GSS 01-APR-2001

DEFINITION 474.dhz98e09.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces unisporus genomic clone 474.dhz98e09.s1, DNA sequence.

ACCESSION A2932397

VERSION A2932397.1 GI:13503310

KEYWORDS GSS.

SOURCE Saccharomyces unisporus.

ORGANISM Saccharomyces unisporus.

REFERENCE 1 (bases 1 to 273)

AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.

TITLE Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis

JOURNAL Unpublished (2001)

COMMENT Contact: Johnston M

Department of Genetics

Washington University Medical School

Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA

Tel: 314 362 2735

Fax: 314 362 7855

Email: mjgenetics.wustl.edu

Class: random plasmid subclone.

FEATURES

Source

1..273

/organism="Saccharomyces unisporus"

/strain="NRRL Y-1556 (CBS 398)"

/db_xref="taxon:27294"

/clone="474.dhz98e09.s1"

/clone_lib="Saccharomyces unisporus NRRL Y-1556"

/note="Random genomic sequence"

BASE COUNT 64 a 78 c 60 g 71 t

ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 273;

Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21

|||||

Db 13 CCGCTTCACTCGCGCTTACTA 33

RESULT 14

LOCUS A2916987 319 bp DNA linear GSS 01-APR-2001

DEFINITION 4911.f663j12.s1 Saccharomyces bayanus MCYC 623-6C Saccharomyces bayanus genomic clone 4911.f663j12.s1, DNA sequence.

ACCESSION A2916987

VERSION A2916987.1 GI:13493511

KEYWORDS GSS.

SOURCE Saccharomyces bayanus.

ORGANISM Saccharomyces bayanus.

REFERENCE 1 (bases 1 to 319)

AUTHORS Clifton,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish,W.R., Waterston,R.H. and Johnston,M.

TITLE Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis

JOURNAL Unpublished (2001)

COMMENT Contact: Johnston M

Department of Genetics

Washington University Medical School

Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA

Tel: 314 362 2735

Fax: 314 362 7855

Email: mjgenetics.wustl.edu

Class: random plasmid subclone.

FEATURES

Location/Qualifiers

1..319

/organism="Saccharomyces bayanus"

/strain="MCYC 623-6C"

/db_xref="taxon:4931"

/clone="4911.f663j12.s1"

/clone_lib="Saccharomyces bayanus MCYC 623-6C"

/note="Random genomic sequence"

BASE COUNT 91 a 57 c 92 g 79 t

ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 319;

Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21

|||||

Db 145 CCGCTTCACTCGCGCTTACTA 125

RESULT 15

LOCUS A2124480 330 bp DNA linear GSS 01-JUN-2000

DEFINITION K025r S. Kluyveri genomic library Saccharomyces kluyveri genomic, DNA sequence.

ACCESSION A2124480

VERSION A2124480.1 GI:8889034

KEYWORDS GSS.

SOURCE Saccharomyces kluyveri.

ORGANISM Saccharomyces kluyveri.

REFERENCE 1 (bases 1 to 330)

AUTHORS Langkjaer,R.B., Nielsen,M.L., Daugaard,P.R., Liu,W. and Piskur,J.

TITLE Yeast chromosomes have been significantly reshaped during their evolutionary history

JOURNAL Unpublished (2000)

COMMENT Contact: Langkjaer RB and Piskur J

Department of Microbiology

Technical University of Denmark

Building 301, DK-2800 Lyngby, Denmark

Tel: 0045 45 45 25 18

Fax: 0045 45 93 28 09

Email: imbl@pop.dtu.dk and imj@pop.dtu.dk

Class: plasmid ends.

FEATURES

Location/Qualifiers

1..330

/organism="Saccharomyces kluyveri"

/strain="NRRL Y-12651r"

/db_xref="taxon:4934"

/clone_lib="S. Kluyveri genomic library"

/note="Vector: pBR322; Partial Sau3A fragments in BamHI"

BASE COUNT 83 a 85 c 69 g 93 t

ORIGIN

Query Match 95.5%; Score 21; DB 12; Length 330;

Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTA 21

|||||

Db 58 CCGCTTCACTCGCGCTTACTA 78

Search completed: October 17, 2002, 10:03:52

Job time : 1161.22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model:

Run on: October 17, 2002, 07:07:41 : Search time 32.7111 Seconds

(without alignments)
172.711 Million cell updates/sec

Title: US-09-780-113D-15

Perfect score: 23

Sequence: 1 accacgactgagcagcactt 23

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/lna/PCYUS.COMB.seq:*

6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.6	72.2	1181	2 US-08-632-598-2	Sequence 2, Appli
2	16.6	72.2	1181	4 US-09-231-240-2	Sequence 2, Appli
3	16.6	72.2	10079	2 US-08-476-866-20	Sequence 20, Appli
4	15.6	67.8	686	4 US-08-943-731-116	Sequence 116, App
5	15.6	67.8	24183	4 US-08-943-731-3	Sequence 3, Appli
6	15.2	66.1	821	4 US-08-735-836C-136	Sequence 136, App
7	15.2	66.1	2246	4 US-09-117-860-17	Sequence 17, Appli
8	15.2	66.1	3435	1 US-08-366-577-1	Sequence 1, Appli
9	15.2	66.1	3435	5 PCT-US96-00005-1	Sequence 1, Appli
10	15.2	66.1	11219	1 US-07-642-734C-1	Sequence 1, Appli
11	15.2	66.1	11219	3 US-08-439-009A-1	Sequence 1, Appli
12	15.2	66.1	51259	3 US-08-781-891-209	Sequence 209, App
13	14.8	64.3	1854	4 US-08-943-731-650	Sequence 650, App
14	14.8	64.3	1854	4 US-09-230-944-9	Sequence 9, Appli
15	14.8	64.3	1874	1 US-08-279-700-3	Sequence 3, Appli
16	14.8	64.3	1919	1 US-08-279-700-5	Sequence 5, Appli
17	14.8	64.3	1919	1 US-08-279-700-7	Sequence 7, Appli
18	14.8	64.3	1919	1 US-08-279-700-11	Sequence 11, Appli
19	14.6	63.5	1778	1 PCT-US91-02958-9	Sequence 9, Appli
20	14.6	63.5	3348	1 US-08-222-616-35	Sequence 35, Appli
21	14.6	63.5	3348	5 PCT-US95-04228-35	Sequence 35, Appli
22	14.6	63.5	5761	1 US-07-749-001-2	Sequence 2, Appli
23	14.6	63.5	5761	1 US-08-154-198-2	Sequence 2, Appli
24	14.6	63.5	5761	1 US-08-463-335-2	Sequence 2, Appli
25	14.6	63.5	5761	2 US-08-464-023A-2	Sequence 2, Appli
26	14.6	63.5	4403765	4 US-09-103-840A-2	Sequence 2, Appli
27	14.2	61.7	880	1 US-08-616-368A-7	Sequence 7, Appli

C 28	14.2	61.7	880	3 US-09-054-298-7	Sequence 7, Appli
C 29	14.2	61.7	880	4 US-08-818-653-7	Sequence 7, Appli
C 30	14.2	61.7	1130	4 US-08-936-165A-248	Sequence 248, App
C 31	14.2	61.7	1652	1 US-08-036-555B-148	Sequence 148, App
C 32	14.2	61.7	1652	1 US-08-469-569-148	Sequence 148, App
C 33	14.2	61.7	1652	1 US-08-249-322A-148	Sequence 148, App
C 34	14.2	61.7	1652	1 US-08-469-526A-148	Sequence 148, App
C 35	14.2	61.7	1652	2 US-08-734-591A-148	Sequence 148, App
C 36	14.2	61.7	1652	2 US-08-469-660-148	Sequence 148, App
C 37	14.2	61.7	1652	3 US-08-341-018-53	Sequence 53, Appli
C 38	14.2	61.7	1652	3 US-08-470-021-148	Sequence 148, App
C 39	14.2	61.7	1652	4 US-08-735-021-148	Sequence 148, App
C 40	14.2	61.7	1652	4 US-08-734-664A-148	Sequence 148, App
C 41	14.2	61.7	1652	4 US-08-470-339-148	Sequence 148, App
C 42	14.2	61.7	1652	5 PCT-US94-05083C-144	Sequence 144, App
C 43	14.2	61.7	1652	5 PCT-US95-06846A-148	Sequence 148, App
C 44	14.2	61.7	16075	3 US-09-096-942-1	Sequence 1, Appli
C 45	14.2	61.7	16075	3 US-09-096-867-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-632-598-2
Sequence 2, Application US/08632598
Patent No. 5886164
GENERAL INFORMATION:
APPLICANT: BIRD, COLIN R
TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA
CORRESPONDENCE ADDRESS: 50
ADDRESS: CUSHMAN DABRY AND CUSHMAN
STREET: 1100 NEW YORK AVENUE N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/632,598
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKUTS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 223355/SEE50112/US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 861-3000
TELEFAX: 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: MOSA
IMMEDIATE SOURCE:
CLONE: EFE GENE
US-08-632-598-2

Query Match 72.2% Score 16.6; DB 2; Length 1181;
Best Local Similarity 82.6% Pred. No. 15;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGACGACGACCTTT 23
||| ||||| ||| ||||| |||
Db 289 ACCTGACTGGAGAAAGCACCCTT 311

RESULT 2

US-09-231-240-2
Sequence 2, Application US/09231240
Patent No. 6262346
GENERAL INFORMATION:
APPLICANT: BIRD, COLIN R
APPLICANT: FLETCHER, JONATHAN D
TITLE OF INVENTION: RIPENING-RELATED GENES FROM BANANA
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DAREY AND CUSHMAN
STREET: 1100 NEW YORK AVENUE N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,240
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,598
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 223355/SSEE50112/US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 861-3000
TELEFAX: 822-0944
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: MUSA
IMMEDIATE SOURCE:
CLONE: EFE GENE
US-09-231-240-2

Query Match 72.2%; Score 16.6; DB 4; Length 1181;
Best Local Similarity 82.6%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGAGCAGCACCCTT 23
||| ||||| ||| ||||| |||
Db 289 ACCTGACTGGAGAAAGCACCCTT 311

RESULT 3

US-08-476-866-20/c
Sequence 20, Application US/08476866
Patent No. 5994339

GENERAL INFORMATION:
APPLICANT: GRABO, JAMES D.
APPLICANT: FRIDOVICH, IRWIN
APPLICANT: OURY, TIM
APPLICANT: DAY, BRIAN J.
APPLICANT: FOLZ, RODNEY J.

APPLICANT: FREEMAN, BRUCE A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE AND MIMETICS THEREOF
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,866
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/322,766
FILING DATE: 13-OCT-1994
APPLICATION NUMBER: US 08/136,207
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-74
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 10079 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 5086..5803
US-08-476-866-20

Query Match 72.2%; Score 16.6; DB 2; Length 10079;
Best Local Similarity 82.6%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ACCACGACTGACGACGACCTTT 23
||| ||||| ||| ||||| |||
Db 1238 ACCTCACTTAGCACCACCTT 1216

RESULT 4

US-08-943-731-116/c
Sequence 116, Application US/08943731
Patent No. 6265157

GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORKKO, JARMO
APPLICANT: ALA-KORKKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:

ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-967-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 686 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-116

Query Match 67.8%; Score 15.6; DB 4; Length 686;
Best Local Similarity 81.8%; Pred. No. 44;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGCAGTGCAGCAGCAGCTTT 23
DB 408 CGAGGACTGAGCAGCAGCTCT 387

RESULT 5
US-08-943-731-3/C
Sequence 3, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISA
APPLICANT: LARSON, ANDREA W.
APPLICANT: PACK, MICHAEL
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA

ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-967-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24183 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-3

Query Match 67.8%; Score 15.6; DB 4; Length 24183;
Best Local Similarity 81.8%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 CCAGCAGTGCAGCAGCAGCTTT 23
DB 4798 CGAGGACTGAGCAGCAGCTCT 4777

RESULT 6
US-08-235-836C-136
Sequence 136, Application US/08235836C
Patent No. 6248562
GENERAL INFORMATION:
APPLICANT: Dunn, John J.
APPLICANT: Luft, Benjamin J.
TITLE OF INVENTION: No. 6248562el Chimeric Proteins Comprising
TITLE OF INVENTION: Borrelia Polypeptides and Uses Therefor
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brookhaven National Laboratory
STREET:
CITY: Upton
STATE: NY
COUNTRY: USA
ZIP: 11973
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,836C
FILING DATE: 29-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,191
FILING DATE: 01-11-93
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Margaret C.

REGISTRATION NUMBER: 25,324
REFERENCE/DOCKET NUMBER: BN193-28A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 282-7338
TELEFAX: (516) 282-3729
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 821 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-235-836C-136

Query Match 66.1%; Score 15.2; DB 4; Length 821;
Best Local Similarity 85.0%; Pred. No. 72;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 1 ACCAGACTGAGCAGCAGC 20
Db 740 ACTAGCTGAGCAGCAGC 759

RESULT 7
US-09-117-860-17
Sequence 17, Application US/09117860A
Patent No. 6338955
GENERAL INFORMATION:
APPLICANT: OGURI, Suguru
APPLICANT: MINOWA, Mari
APPLICANT: YOSHIDA, Aruto
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: TAKEUCHI, Makoto
TITLE OF INVENTION: NOVEL 1-4-N-ACETYLGLUCOSAMINYLTRANSFERASE AND GENE
TITLE OF INVENTION: ENCODING THE SAME
FILE REFERENCE: 081356/0119
CURRENT APPLICATION NUMBER: US/09/117,860A
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: WO PCT/JP97/04546
EARLIER FILING DATE: 1997-12-10
EARLIER APPLICATION NUMBER: JP 161462/1997
EARLIER FILING DATE: 1997-06-18
EARLIER APPLICATION NUMBER: JP 332411/1996
EARLIER FILING DATE: 1996-12-12
NUMBER OF SEQ ID NOS: 63
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 17
LENGTH: 2246
TYPE: DNA
ORGANISM: Bovine
FEATURE:
NAME/KEY: CDS
LOCATION: (288)..(1892)
US-09-117-860-17

Query Match 66.1%; Score 15.2; DB 4; Length 2246;
Best Local Similarity 85.0%; Pred. No. 83;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 4 AGAGCTGAGCAGCAGCCTT 23
Db 1950 ACAGCTGAGCAGCAGCTT 1969

RESULT 8
US-08-366-577-1/c
Sequence 1, Application US/08366577
Patent No. 5728523
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
TITLE OF INVENTION: POLYMERASE DELTA MUTATIONS IN COLORECTAL
TITLE OF INVENTION: TUMORS WITH REPLICATION ERRORS
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/366,577
FILING DATE: 12-DEC-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107,48554
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBWB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3435 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 43..3364
US-08-366-577-1

Query Match 66.1%; Score 15.2; DB 1; Length 3435;
Best Local Similarity 85.0%; Pred. No. 88;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

UY 2 CCAGACTGAGCAGCAGCT 21
Db 1691 CCAGACTGAGCAGCTACT 1672

RESULT 9
PCT-US96-00005-1/c
Sequence 1, Application PC/T059600005
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
TITLE OF INVENTION: POLYMERASE DELTA MUTATIONS IN COLORECTAL
TITLE OF INVENTION: TUMORS WITH REPLICATION ERRORS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 1001 G Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00005
FILING DATE: 2-JAN-96


```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.53505
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3435 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 43..3364
PCT-US96-00005-1

Query Match 66.18; Score 15.2; DB 5; Length 3435;
Best Local Similarity 85.08; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCAGAGCTGACGACGACCT 21
    |||||
DB 1691 CCAAGACTGACGACGACT 1672

RESULT 10
US-07-642-734C-1/C
; Sequence 1, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:
; APPLICANT: Katz, L
; APPLICANT: Donadio, S
; TITLE OF INVENTION: Recombinant DNA Method for Producing
; TITLE OF INVENTION: Erythromycin Analogs
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman
; STREET: Abbott Laboratories D377/AF6D-2 One Abbott
; STREET: Park Rd
; CITY: Abbott Park
; STATE: IL
; COUNTRY: US
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,734C
; FILING DATE: 17-JAN-91
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Danckers, Andreas M
; REGISTRATION NUMBER: 32652
; REFERENCE/DOCKET NUMBER: 4952.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9396
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11219 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 2338
FEATURE:
NAME/KEY: misc_feature
LOCATION: 744..6659
OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
OTHER INFORMATION: MODULE 1"
OTHER INFORMATION: /label= FUNCTION
FEATURE:
NAME/KEY: CDS
LOCATION: 744..11219
OTHER INFORMATION: /function= "gene- "erya""
OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
OTHER INFORMATION: 6-deoxyerythronolide B""
FEATURE:
NAME/KEY: misc_feature
LOCATION: 744..1868
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain 1 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1998..2198
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 1 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2250..3626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3831..4811
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase 2 domain of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5574..6125
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6369..6626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 2 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..11219
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..8066
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8262..9305
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9906..10454
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10707..10964
```

OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 2"
US-07-642-734C-1

Query Match 65.2%, Score 15; DB 1; Length 11219;
Best Local Similarity 78.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTT 23
||||| ||||| ||||| ||||| |||||

Db 4333 ACCACGACGACGACGACGCTT 4311

RESULT 11
US-08-439-009A-1/C

Sequence 1, Application US/08439009A
Patent No. 6004787
GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
SPECIFIC POLYPEPTIDES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/AB6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952-US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11219 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 2338

FEATURE:
NAME/KEY: misc_feature
LOCATION: 744..6659
OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
OTHER INFORMATION: MODULE 1"
OTHER INFORMATION: /label= FUNCTION
FEATURE:
NAME/KEY: CDS
LOCATION: 744..11219
OTHER INFORMATION: /function= "gene= "eryA"
OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature

LOCATION: 744..1868
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain 1 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1998..2198
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 1 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2250..3626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3831..4811
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase 2 domain of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5574..6125
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6369..6626
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain 2 of module 1"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..11219
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6678..8066
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8262..9305
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9906..10454
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 2"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10707..10964
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 2"
US-08-439-009A-1

Query Match 65.2%, Score 15; DB 3; Length 11219;
Best Local Similarity 78.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTT 23
||||| ||||| ||||| ||||| |||||

Db 4333 ACCACGACGACGACGACGCTT 4311

RESULT 12
US-08-781-891-209

Sequence 209, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620lemburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 65.2%; Score 15; DB 3; Length 51259;
Best Local Similarity 78.3%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACCAGGACTGACGACGACCTTT 23
Db 49944 ACCATGGCTGACGATGACATCTCT 49966

RESULT 13
US-08-943-731-650
Sequence 650, Application US/08943731
Patent No. 6265157
GENERAL INFORMATION:
APPLICANT: PROCKOP, DARWIN J.
APPLICANT: SPOTILA, LORETTA D.
APPLICANT: DELTAS, CONSTANTINOS D.
APPLICANT: SEREDA, LARISSA
APPLICANT: LARSON, ANDREA W.
APPLICANT: COLIGE, ALAIN
APPLICANT: EARLY, JAMES
APPLICANT: KORRKO, JARMO
APPLICANT: ALA-KORRKO, LEENA, et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
NUMBER OF SEQUENCES: 666
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
STREET: FLR
CITY: PHILADELPHIA
STATE: PA
COUNTRY: USA
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,731
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/212,322
FILING DATE: 14-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/803,628
FILING DATE: 03-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DOYLE LEARY Ph.D., KATHRYN
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9598-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 650:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-943-731-650

Query Match 64.3%; Score 14.8; DB 4; Length 20;
Best Local Similarity 88.9%; Pred. No. 68;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GACTGACGACGACGACCTTT 23
Db 2 CAGTACGACGACGACCTCT 19

RESULT 14
US-09-230-944-9
Sequence 9, Application US/09230944
Patent No. 6277380
GENERAL INFORMATION:
APPLICANT: UEDA, Shigeharu
APPLICANT: WATANABE, Michiko
APPLICANT: KAWANISHI, Hitomi
TITLE OF INVENTION: Measles Virus Mutant Antigen and Gene Coding for the
FILE REFERENCE: 0216-0407P
CURRENT APPLICATION NUMBER: US/09/230,944
CURRENT FILING DATE: 1999-02-04
EARLIER APPLICATION NUMBER: PCT/JP98/02481
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 9-164285 JAPAN
EARLIER FILING DATE: 1997-06-04
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 1854
TYPE: DNA
ORGANISM: Measles virus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1851)
FEATURE:
OTHER INFORMATION: Attenuated measles virus NA strain
FEATURE:
OTHER INFORMATION: any n or Xaa = Unknown
US-09-230-944-9

Query Match 64.3%; Score 14.8; DB 4; Length 1854;
Best Local Similarity 88.9%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CACGACTGAGCAGCACC 20
 Db 740 CACACTGAGCAGCACC 757

RESULT 15

US-08-279-700-3
 ; Sequence 3, Application US/08279700
 ; Patent No. 5578448
 ; GENERAL INFORMATION:
 ; APPLICANT: ROTA, Jennifer S.
 ; APPLICANT: BELINI, William J.
 ; TITLE OF INVENTION: WILD-TYPE MEASLES VIRUS GLYCOPROTEINS:
 ; TITLE OF INVENTION: VACCINE AND DETECTION METHOD THEREFOR
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentia Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/279,700
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/866,033A
 ; FILING DATE: 19920408
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 40399/140 NIHD
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ. ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1874 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 21..1874
 ; US-08-279-700-3

Query Match 64.3%; Score 14.8; DB 1; Length 1874;
 Best Local Similarity 88.9%; Pred. No. 1.3e+02;
 Matches 16; Conservative 0; Mismatches 2; Indels 0;

OY 3 CACGACTGAGCAGCACC 20
 Db 760 CACACTGAGCAGCACC 777

Search completed: October 17, 2002, 08:52:21
 Job time : 65.7111 secs

AUTHORS Brown,S., Rouse,G., Hutchings,P. and Colgan,D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6
 College Street, Sydney, NSW 2000, Australia

FEATURES
 Source Location/Qualifiers

1.307
 /organism="Mesochaelopterus sp. AMW22402"
 /specimen_voucher="AMW22402"
 /db_xref="taxon:104707"
 <1..>307
 /note="D1 expansion region"
 /product="28S ribosomal RNA"

BASE COUNT 73 a 78 c 90 g 66 t
 ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTCGCGCTTACTAG 22
 |||||
 Db 60 CCGCTTCACTCGCGCTTACTAG 39

Search completed: October 17, 2002, 09:24:05
 Job time : 939.733 secs

Db 58 CCGCTTACTCGCGCTTACTAG 37
 |||
 RESULT 12
 AF185171/c 305 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Polyophthalmus pictus 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185171
 VERSION AF185171.1 GI:6013352
 KEYWORDS
 SOURCE Polyophthalmus pictus.
 ORGANISM Polyophthalmus pictus.
 Eukaryota; Metazoa; Annelida; Polychaeta; Opheliida; Opheliidae;
 Polyophthalmus.
 REFERENCE 1 (bases 1 to 305)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 305)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 FEATURES
 SOURCE
 1.307
 /organism="Polyophthalmus pictus"
 /specimen_voucher="AMW10095"
 /db_xref="taxon:104727"
 <1..>305
 /note="D1 expansion region"
 /product="28S ribosomal RNA"
 BASE COUNT 61 a 84 c 97 g 59 t 4 others
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 305;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 |||
 RESULT 13
 AF185150/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Amphiglena terebro 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185150
 VERSION AF185150.1 GI:6013331
 KEYWORDS
 SOURCE Amphiglena terebro.
 ORGANISM Amphiglena terebro.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliipalpata;
 Sabellida; Sabellidae; Amphiglena.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 FEATURES
 SOURCE
 1.307
 /organism="Amphiglena terebro"
 /specimen_voucher="AMW21360"
 /db_xref="taxon:104683"
 <1..>307
 /note="D1 expansion region"

BASE COUNT 69 a 79 c 92 g 65 t 2 others
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 |||
 RESULT 14
 AF185154/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Glycera tridactyla 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185154
 VERSION AF185154.2 GI:6073854
 KEYWORDS
 SOURCE Glycera tridactyla.
 ORGANISM Glycera tridactyla.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
 Phyllodoctida; Glyceridae; Glycera.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1999) Evolutionary Biology, Australian Museum, 6 College Street, Sydney, NSW 2000, Australia
 COMMENT On Oct 20, 1999 this sequence version replaced gi:6013335.
 FEATURES
 SOURCE
 1.307
 /organism="Glycera tridactyla"
 /specimen_voucher="AMW196835"
 /db_xref="taxon:1047710"
 <1..>307
 /note="D1 expansion region"
 /product="28S ribosomal RNA"
 BASE COUNT 62 a 81 c 97 g 67 t
 ORIGIN
 Query Match 100.0%; Score 22; DB 3; Length 307;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 CCGCTTACTCGCGCTTACTAG 22
 |||
 Db 60 CCGCTTACTCGCGCTTACTAG 39
 |||
 RESULT 15
 AF185163/c 307 bp DNA linear INV 20-OCT-1999
 LOCUS
 DEFINITION Mesochaetopterus sp. AMW22402 28S ribosomal RNA gene, partial sequence.
 ACCESSION AF185163
 VERSION AF185163.1 GI:6013344
 KEYWORDS
 SOURCE Mesochaetopterus sp. AMW22402.
 ORGANISM Mesochaetopterus sp. AMW22402.
 Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canaliipalpata;
 Spionida; Chaetopteridae; Mesochaetopteridae.
 REFERENCE 1 (bases 1 to 307)
 AUTHORS Brown, S., Rouse, G., Hutchings, P. and Colgan, D.
 TITLE Assessing the usefulness of histone H3, U2 snRNA and 28S rDNA in analyses of polychaete relationships
 JOURNAL Aust. J. Zool. 47 (1999) In press
 REFERENCE 2 (bases 1 to 307)

FEATURES
source location/Qualifiers
1.283
/organism="uncultured basidiomycete Cons3"
/isolate="Cons3"
/db_xref="taxon:150676"
/country="France: Antifer Harbour"
<1..>283
/product="24S large subunit ribosomal RNA"
78 a 53 c 89 g 63 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 22; DB 8; Length 283;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTAG 22
|||||
DB 46 CCGCTTCACTCGCGCTTACTAG 25

RESULT 9
AF339505/c 295 bp DNA linear INV 04-JUN-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Aplysia depilans.
Aplysia depilans.
Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia;
Anaspidae; Aplysiidae; Aplysia.
1 (bases 1 to 295)
Tillier,S., Masselot,M., Guerdoux,J. and Tillier,A.
Monophyly of major Gastropod taxa tested from partial 28S rRNA
sequences, with emphasis on Euthyrea and hot vent limpets
Peltospiridae
Nautilus 108 (supplement 2), 122-140 (1994)
2 (bases 1 to 295)
Dayrat,B., Tillier,A., Lecointre,G. and Tillier,S.
New clades of euthyreaean gastropods (Mollusca) from 28S rRNA
sequences
Molecular phylogenetics and evolution. 19 (2), 225-235 (2001)
3 (bases 1 to 295)
11341805
Tillier,S., Masselot,M., Guerdoux,J. and Tillier,A.
Direct Submission
Submitted (23-JAN-2001) Laboratoire de Biologie des Invertébrés
Marins et Malacologie, Institut de Systematique FR 1541, Muséum
National d'Histoire Naturelle, 43 rue Cuvier, Paris 75005, France

FEATURES
source location/Qualifiers
1..295
/organism="Aplysia depilans"
/db_xref="taxon:76186"
<1..>295
/note="DI domain"
/product="28S ribosomal RNA"
72 a 76 c 91 g 55 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 295;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTAG 22
|||||
DB 100 CCGCTTCACTCGCGCTTACTAG 79

RESULT 10
AF265622/c 303 bp DNA linear INV 09-JUL-2000

LOCUS
AF265622

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Madracis pharensis 28S ribosomal RNA gene, partial sequence.
AF265622
AF265622.1 GI:8980711
Madracis pharensis.
Madracis pharensis.
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoeniina; Pocilloporidae; Madracis.
1 (bases 1 to 303)
Romano,S.L. and Cairns,S.D.
Molecular phylogenetic hypotheses for the evolution of
scleractinian corals
unpublished
2 (bases 1 to 303)
Romano,S.L. and Cairns,S.D.
Direct Submission
Submitted (09-MAY-2000) Laboratory of Molecular Systematics,
National Museum of Natural History, Smithsonian Institution,
Washington, DC 20560-0163, USA

FEATURES
source location/Qualifiers
1..303
/organism="Madracis pharensis"
/db_xref="taxon:123772"
<1..>303
/product="28S ribosomal RNA"
83 a 64 c 92 g 64 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 303;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTAG 22
|||||
DB 58 CCGCTTCACTCGCGCTTACTAG 37

RESULT 11
AF265642/c 304 bp DNA linear INV 09-JUL-2000

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Caryophyllia inornata.
Caryophyllia inornata.
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Caryophyllina; Caryophylliidae; Caryophyllia.
1 (bases 1 to 304)
Romano,S.L. and Cairns,S.D.
Molecular phylogenetic hypotheses for the evolution of
scleractinian corals
unpublished
2 (bases 1 to 304)
Romano,S.L. and Cairns,S.D.
Direct Submission
Submitted (09-MAY-2000) Laboratory of Molecular Systematics,
National Museum of Natural History, Smithsonian Institution,
Washington, DC 20560-0163, USA

FEATURES
source location/Qualifiers
1..304
/organism="Caryophyllia inornata"
/db_xref="taxon:130053"
<1..>304
/product="28S ribosomal RNA"
80 a 68 c 92 g 64 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 304;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCGCTTCACTCGCGCTTACTAG 22

```

VERSION      065520.1  GI:1553147
KEYWORDS
SOURCE
ORGANISM      Seriatopora hystrix.
              Seriatopora hystrix
              Eukaryota; Metazoa; Anthozoa; Zoantharia; Scleractinia;
              Asterozoa; Pocilloporidae; Seriatopora.
REFERENCE
AUTHORS      1 (bases 1 to 222)
              Chen, C.A., Odorico, D.M., ten Lohuis, M., Veron, J.E. and Miller, D.J.
              Systematic relationships within the Anthozoa (Cnidaria: Anthozoa)
              using the 5'-end of the 28S rDNA
              Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
JOURNAL
MEDLINE
REFERENCE
AUTHORS      2 (bases 1 to 222)
              Veron, J.E.N., Odorico, D.M., Chen, C.A. and Miller, D.J.
              Reassessing evolutionary relationships of scleractinian corals
              Coral Reefs 15, 1-9 (1996)
REFERENCE
AUTHORS      3 (bases 1 to 222)
              Odorico, D.M., Chen, C.A. and Miller, D.J.
              Direct Submission
JOURNAL
TITLE
SOURCE
FEATURES
source
1..222
/organism="Seriatopora hystrix"
/db_xref="taxon:51070"
1..>222
/product="28S ribosomal RNA"
BASE COUNT   59 a      53 c      68 g      42 t
ORIGIN
Query Match      100.0%; Score 22; DB 3; Length 222;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CCGCTTCACTGCGCGTTACTAG 22
         |||
Db      58  CCGCTTCACTGCGCGTTACTAG 37

RESULT 6
TS28S259/c
LOCUS      TS28S259      259 bp      mRNA      linear      INV 20-AUG-1996
DEFINITION Tubifex sp. gene for 28S ribosomal RNA (partial; 259 bp).
ACCESSION  X90690
VERSION     X90690.1  GI:1067238
KEYWORDS   28S ribosomal RNA; 28S rRNA gene.
SOURCE
ORGANISM   Tubifex sp.
            Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
            Tubificina; Tubificidae; Tubifex.
REFERENCE
AUTHORS     1 (bases 1 to 259)
            Friedrich, M.
            Direct Submission
            Submitted (09-AUG-1995) M. Friedrich, Zoological Institute, Univ.
            Muenchen, Luisenstr. 14, D-80333 Muenchen, FRG
JOURNAL
TITLE       2 (bases 1 to 259)
            Friedrich, M. and Tautz, D.
            Ribosomal DNA phylogeny of the major extant arthropod classes and
            the evolution of mytilopods
            Nature 376 (6536), 165-167 (1995)
JOURNAL
MEDLINE
FEATURES
source
Location/Qualifiers
1..259
/organism="Tubifex sp."
/db_xref="taxon:44784"
/dev_stage="adult"
<1..>259
/gene="28S rRNA"
/product="28S ribosomal RNA"
1..259
/gene="28S rRNA"
BASE COUNT   54 a      75 c      84 g      46 t
ORIGIN

```

```

Query Match      100.0%; Score 22; DB 3; Length 259;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CCGCTTCACTGCGCGTTACTAG 22
         |||
Db      30  CCGCTTCACTGCGCGTTACTAG 9

RESULT 7
TSU75861/c
LOCUS      TSU75861      276 bp      RNA      linear      INV 23-DEC-1996
DEFINITION Theodoxus sp. 28S ribosomal RNA gene, partial sequence.
ACCESSION  U75861
VERSION     U75861.1  GI:1750151
KEYWORDS
SOURCE
ORGANISM     Theodoxus sp.
            Theodoxus sp.
            Eukaryota; Metazoa; Mollusca; Gastropoda; Neritimorpha; Neritimorpha;
            Theodoxus.
            Theodoxus.
            1 (bases 1 to 276)
            McArthur, A.G.
            Molecular investigation of the evolutionary origins of hydrochrymal
            vent gastropods
            Thesis, University of Victoria, Canada. (1996)
JOURNAL
REFERENCE    2 (bases 1 to 276)
            McArthur, A.G.
            Direct Submission
            Submitted (23-OCT-1996) Biology, University of Victoria, P.O. Box
            3020, Victoria, BC V8W 3N5, Canada
JOURNAL
TITLE
SOURCE
FEATURES
source
1..276
/organism="Theodoxus sp."
/db_xref="taxon:55034"
/dev_stage="adult"
<1..>276
/product="28S ribosomal RNA gene"
BASE COUNT   61 a      76 c      91 g      48 t
ORIGIN
Query Match      100.0%; Score 22; DB 3; Length 276;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CCGCTTCACTGCGCGTTACTAG 22
         |||
Db      61  CCGCTTCACTGCGCGTTACTAG 40

RESULT 8
AF318260/c
LOCUS      AF318260      283 bp      DNA      linear      PLN 06-MAR-2001
DEFINITION uncultured basidiomycete Cons3 24S large subunit ribosomal RNA
            gene, partial sequence.
ACCESSION  AF318260
VERSION     AF318260.1  GI:13236258
KEYWORDS
SOURCE
ORGANISM     uncultured basidiomycete Cons3.
            uncultured basidiomycete Cons3
            Eukaryota; Fungi; Basidiomycota; environmental samples.
REFERENCE
AUTHORS     1 (bases 1 to 283)
            Guillon, P., Nezan, E., Cueff, V., Erard-Le Denn, E., Cambon, M.-A.,
            Gentien, P. and Barbier, G.
            Semi-nested PCR detection of three toxic dinoflagellate genera
            (Alexandrium, Dinophysis, and Gymnodinium) in sea water column and
            sediment from French coasts
            Unpublished
JOURNAL
REFERENCE    2 (bases 1 to 283)
            Guillon, P., Nezan, E., Cueff, V., Erard-Le Denn, E., Cambon, M.-A.,
            Gentien, P. and Barbier, G.
            Direct Submission
            Submitted (02-NOV-2000) DRV VP CMM, IFREMER, B.P. 70, Prest 29240,

```


ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 218;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 59 CCGCTTCACTGCGCGTTACTAG 38

RESULT 2
E64938/c
LOCUS
DEFINITION
cdna sequence of gene participating in induction of resistivity in
plant.
ACCESSION
E64938
VERSION
E64938.1 GI:13017969
KEYWORDS
JP 1999318477-A/70.
SOURCE
Nicotiana tabacum
ORGANISM
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE
1 (bases 1 to 218)
AUTHORS
Peter/H.S.G.C. and Gregor/I.S.K.K.
TITLE
cdna sequence of gene participating in induction of resistivity in
JOURNAL
Bayer AG
Parent: JP 1999318477-A 70 24-NOV-1999;

COMMENT
OS Nicotiana tabacum
PN JP 1999318477-A/70
PD 24-NOV-1999
PF 1999075762
PR 19-MAR-1999 JP 1999075762
PI 25-MAR-1998 DE 19813048.1
SI PETER HELMUTO SCHREIER,CHRISTINE GARBERS,GREGOR LANGEN, PI
SIEGRIED KIDOROUSKI
PC C12N15/09,A01H5/00,A01N65/02,C07K14/415,C07K16/16,C12N1/15, PC
C12N1/21,
PC C12N5/10,C12N5/10,C12P21/02,C12Q1/68,G01N33/48,G01N33/566// PC
(C12N1/21,C12R1:01),(C12P21/02,C12R1:91),C12N15/00,C12N5/00, PC
C12N5/00
CC
FH Key
FT source
FT Location/Qualifiers
1..218
/organism="Nicotiana tabacum"
/db_xref="taxon:4097"
BASE COUNT 68 a 43 c 65 g 42 t
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 218;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 59 CCGCTTCACTGCGCGTTACTAG 38

RESULT 3
CS069678/c
LOCUS
DEFINITION
Cerianthus sp. 28S ribosomal RNA gene, partial sequence.
ACCESSION
U69678
VERSION
U69678.1 GI:1617542
KEYWORDS
KRYWIDOS
SOURCE
Cerianthus sp.
ORGANISM
Cerianthus sp.
Eukaryota; Metazoa; Chnidaria; Anthozoa; Ceriantipatharia;
Ceriantbaria; Ceriantbaridae; Ceriantbus.
REFERENCE
1 (bases 1 to 219)

AUTHORS
TITLE
Chen,C.A., Odorico,D.M., ten-Johnis,M., Veron,J.E. and Miller,D.J.
Systematic relationships within the Anthozoa (Chnidaria: Anthozoa)
using the 5'-end of the 28S rDNA
JOURNAL
Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
MEDLINE
95392827
REFERENCE
2 (bases 1 to 219)
AUTHORS
Chen,C.A., Odorico,D.M., ten-Johnis,M., Veron,J.E. and Miller,D.J.
TITLE
Direct Submission
JOURNAL
Submitted (06-SEP-1996) Biochemistry and Molecular Biology, James
Cook University, Townsville, Queensland 4811, Australia
Location/Qualifiers
1..219
/organism="Cerianthus sp."
/db_xref="taxon:51771"
1..>219
/product="28S ribosomal RNA"
BASE COUNT 61 a 48 c 59 g 51 t
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 219;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 58 CCGCTTCACTGCGCGTTACTAG 37

RESULT 4
CAU69690/c
LOCUS
DEFINITION
Corynactis australis 28S ribosomal RNA gene, partial sequence.
ACCESSION
U69690
VERSION
U69690.1 GI:1617541
KEYWORDS
SOURCE
Corynactis australis.
ORGANISM
Corynactis australis.
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Corallimorphidae; Corynactis.
REFERENCE
1 (bases 1 to 220)
AUTHORS
Chen,C.A., Odorico,D.M., ten-Johnis,M., Veron,J.E. and Miller,D.J.
TITLE
Systematic relationships within the Anthozoa (Chnidaria: Anthozoa)
using the 5'-end of the 28S rDNA
JOURNAL
Mol. Phylogenet. Evol. 4 (2), 175-183 (1995)
MEDLINE
95392827
REFERENCE
2 (bases 1 to 220)
AUTHORS
Chen,C.A., Odorico,D.M., ten-Johnis,M., Veron,J.E. and Miller,D.J.
TITLE
Direct Submission
JOURNAL
Submitted (06-SEP-1996) Biochemistry and Molecular Biology, James
Cook University, Townsville, Queensland 4811, Australia
Location/Qualifiers
1..220
/organism="Corynactis australis"
/db_xref="taxon:51772"
1..>220
/product="28S ribosomal RNA"
BASE COUNT 54 a 55 c 67 g 44 t
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 220;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTTCACTGCGCGTTACTAG 22
|||||
Db 58 CCGCTTCACTGCGCGTTACTAG 37

RESULT 5
SH065520/c
LOCUS
DEFINITION
Seriatopora hystrix 28S ribosomal RNA gene, partial sequence.
ACCESSION
U65520

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:07:35 : Search time 935.733 seconds

(without alignments)
492.003 Million cell updates/sec

Title: US-09-780-113D-8

Perfect score: 22

Sequence: 1 ccgcctcactgcgcgttactag 22

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenBankl:*
1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_sys:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hig_hum:*
31: em_hig_inv:*
32: em_hig_other:*
33: em_hig_inv:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description

c	1	22	100.0	218	6	AX014514	AX014514 Sequence
c	2	22	100.0	218	6	E64938	E64938 cDNA sequen
c	3	22	100.0	219	3	CS069678	U69678 Cerianthus
c	4	22	100.0	220	3	CA069690	U69690 Corynactis
c	5	22	100.0	222	3	SH065520	U65520 Seriatopora
c	6	22	100.0	229	3	TS078529	X90690 Tufflex sp.
c	7	22	100.0	276	3	TS078561	U75861 Theodorus s
c	8	22	100.0	283	8	AF181826	AF181826 Unculture
c	9	22	100.0	285	3	AF181826	AF181826 Unculture
c	10	22	100.0	303	3	AF181826	AF181826 Unculture
c	11	22	100.0	303	3	AF181826	AF181826 Unculture
c	12	22	100.0	304	3	AF181826	AF181826 Unculture
c	13	22	100.0	305	3	AF181826	AF181826 Unculture
c	14	22	100.0	307	3	AF181826	AF181826 Unculture
c	15	22	100.0	307	3	AF181826	AF181826 Unculture
c	16	22	100.0	307	3	AF181826	AF181826 Unculture
c	17	22	100.0	310	3	AF181826	AF181826 Unculture
c	18	22	100.0	311	3	AF181826	AF181826 Unculture
c	19	22	100.0	317	8	AF181826	AF181826 Unculture
c	20	22	100.0	323	8	AF181826	AF181826 Unculture
c	21	22	100.0	343	3	ORDNA28S	AF326085 Phaloccep
c	22	22	100.0	347	3	MSPDNA28	AF326085 Phaloccep
c	23	22	100.0	347	3	PADNA28R	X80651 O. rathkei D
c	24	22	100.0	349	3	MSPDNA28R	X80648 Mesoneilla
c	25	22	100.0	360	3	AF358798	X80647 M. palmaria P
c	26	22	100.0	370	3	AF358815	AF358815 Chyllina f
c	27	22	100.0	370	3	LO28RRN5	X57259 L. octona s
c	28	22	100.0	371	3	PMJ225826	AJ225826 Petrolion
c	29	22	100.0	374	3	CCJ225827	U61652 Fusarium co
c	30	22	100.0	417	8	FCU61652	U61652 Fusarium co
c	31	22	100.0	452	8	MTU43478	U43478 Massaria in
c	32	22	100.0	456	8	LDU43474	U43474 Leptosphaer
c	33	22	100.0	457	8	CDU43481	U43481 Clathrospor
c	34	22	100.0	457	8	LDU43473	U43473 Leptosphaer
c	35	22	100.0	457	8	LDU43475	U43475 Leptosphaer
c	36	22	100.0	458	8	AF300719	AF300719 Salai roo
c	37	22	100.0	458	8	OFU43472	U43472 Ophiobolus
c	38	22	100.0	458	8	OHU43471	U43471 Ophiobolus
c	39	22	100.0	459	8	PTU43477	U43477 Pyrenophora
c	40	22	100.0	488	8	AF158614	AF158614 Fusarium
c	41	22	100.0	492	8	AF261555	AF261555 Phaeococc
c	42	22	100.0	518	8	PN1276065	AF300718 Salai roo
c	43	22	100.0	520	8	AF300718	U45628 Candida hal
c	44	22	100.0	521	8	CHU45828	U45830 Candida man
c	45	22	100.0	521	8	CHU45830	

ALIGNMENTS

RESULT 1
LOCUS AX014514/c
DEFINITION Sequence 70 from Patent EP0953640.
ACCESSION AX014514
VERSION AX014514.1 GI:10040789
KEYWORDS
SOURCE
ORGANISM
Common tobacco.
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.

REFERENCE
AUTHORS Kiedrowski, S.D., Gardner, C.D., Jansen, G.D. and Schreier, P.H.
TITLE 1 (bases 1 to 218)
JOURNAL Cdna sequences of genes involved in the induction of resistance in
plants
Patent: EP 0953640-A 70 03-NOV-1999;
BAYER AG (DE)

FEATURES
source
location/Qualifiers
1..218
/organism="Nicotiana tabacum"
/db_xref="taxon:4097"
BASE COUNT 68 a 43 c 65 g 42 t

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2002, 08:07:35 ; Search time 978.267 Seconds
(without alignments)
492.003 Million cell updates/sec

Title: US-09-780-113D-15

Perfect score: 23

Sequence: 1 accacgactgagcagcaccctt 23

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.com:*
5: gb.ov:*
6: gb.pat:*
7: gb.pb:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vi:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htg.in:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	712	8	AF042820
2	23	100.0	713	8	AF086948
3	23	100.0	715	8	AF210741
4	23	100.0	715	8	AF211256
5	23	100.0	4103	8	AF409124
6	18.2	79.1	126149	9	AC006352
7	18.2	79.1	169710	2	AL627202
8	18.2	79.1	188930	2	AL592114
9	18.2	79.1	197083	2	AC091106
10	18.2	79.1	214025	2	AC007882
11	18.2	79.1	227856	2	AC007908
12	18.2	79.1	234542	2	HUAC002041
13	17.8	77.4	215359	9	AC012495
14	17.4	75.7	34105	2	CNS000YVE
15	17.4	75.7	197764	2	AC021605
16	17.2	74.8	26689	9	AL499629
17	17.2	74.8	43772	9	AC004791
18	17.2	74.8	46708	2	AL359434
19	17.2	74.8	73806	2	AC095461
20	17.2	74.8	73108	2	AC005363
21	17.2	74.8	80668	9	AC005606
22	17.2	74.8	81579	9	AE006640
23	17.2	74.8	107549	9	AC074178
24	17.2	74.8	123331	9	AC005164
25	17.2	74.8	128628	9	AC093758
26	17.2	74.8	163031	9	AC024947
27	17.2	74.8	163889	2	AC004958
28	17.2	74.8	175996	9	AF069291
29	17.2	74.8	184919	9	AC010539
30	17.2	74.8	207962	2	AC015679
31	17.2	74.8	238169	2	AC091396
32	17.2	74.8	242965	2	AC023175
33	17.2	74.8	320250	2	AF117829
34	17.2	73.9	154921	2	AC109793
35	17.2	73.9	205140	2	AC109795
36	17.2	73.9	335372	2	AC094935
37	16.8	73.0	280	9	HSCRP06
38	16.8	73.0	355	13	AF349055
39	16.8	73.0	42856	2	AC100075
40	16.8	73.0	110000	2	AC105915_0
41	16.8	73.0	110000	2	AC005079_2
42	16.8	73.0	121459	2	AL137860
43	16.8	73.0	168042	2	AC022529
44	16.8	73.0	170128	2	AC005059
45	16.8	73.0	175390	2	AC094641

ALIGNMENTS

RESULT 1
LOCUS AF042820/c 712 bp DNA linear PL# 01-FEB-1998
DEFINITION Heterosigma akashiwo 24S large subunit ribosomal RNA gene, partial sequence.
ACCESSION AF042820
VERSION AF042820.1 GI:2827390
KEYWORDS
SOURCE
ORGANISM Heterosigma akashiwo.
Heterosigma akashiwo.
Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
REFERENCE
1 (bases 1 to 712)
Lee, S.W., Park, C.G. and Park, Y.S.
24S ribosomal RNA sequence analysis of dinoflagellates isolated from red-tide in southern coast of Korea
TITLE
Unpublished
JOURNAL
REFERENCE
1 (bases 1 to 712)
Lee, S.W., Park, C.G. and Park, Y.S.
Direct Submission
JOURNAL
Submitted (14-JAN-1998) Department of Microbiology, Inje University, Ochangdong 607, Kimae 621-749, Korea
FEATURES
source
1. 712

/organism="Heterosigma akashiwo"
/db_xref="taxon:2829"
/note="Isolated from red-tide sea water in Masan Bay, Korea; classified by morphological studies"
<1..>712
/note="variable domains D1 and D2"
/product="24S large subunit ribosomal RNA"
BASE COUNT 164 a 149 c 211 g 188 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 712;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTTT 23
|||||
Db 149 ACCACGACTGACGACGACCTTT 127

RESULT 2
AF086948/c 713 bp DNA linear PLN 15-OCT-1998
LOCUS
DEFINITION Heterosigma akashiwo large subunit ribosomal RNA gene, partial
sequence.
AF086948
VERSION AF086948.1 GI:3695263
KEYWORDS
SOURCE Heterosigma akashiwo.
ORGANISM Heterosigma akashiwo
REFERENCE 1 (bases 1 to 713)
Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
AUTHORS Connell, L.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) Environmental Conservation Division,
National Marine Fisheries Service, 2725 Montlake Blvd East,
Seattle, WA 98112, USA
FEATURES
source 1..713
location/Qualifiers
1..713
/organism="Heterosigma akashiwo"
/strain="CCMP-452"
/db_xref="taxon:2829"
<1..>713
/note="contains variable domains D1-D3"
/product="large subunit ribosomal RNA"
BASE COUNT 165 a 148 c 210 g 190 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 713;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCACGACTGACGACGACCTTT 23
|||||
Db 147 ACCACGACTGACGACGACCTTT 125

RESULT 3
AF210741/c 715 bp DNA linear "N 02-JAN-2001
LOCUS
DEFINITION Heterosigma akashiwo large subunit ribosomal RNA gene, partial
sequence.
AF210741
VERSION AF210741.1 GI:12003313
KEYWORDS
SOURCE Heterosigma akashiwo.
ORGANISM Heterosigma akashiwo
Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
REFERENCE 1 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Schollin, C.A., Bergquist, P.R., and Bergquist, P.L.
TITLE Detection and Enumeration of Heterosigma akashiwo and Fibrocapsa
japonica (Raphidophyceae) using rRNA-targeted Oligonucleotide
Probes
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Phylogeny of the Raphidophytes Based on Large-Subunit rRNA Gene
Sequences
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1999) Research and Development, Monterey Bay
Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA
95039-0628, USA
FEATURES
source location/Qualifiers
1..715
/organism="Heterosigma akashiwo"
/strain="CAMR10"
/db_xref="taxon:147349"
/note="similar to Heterosigma akashiwo sequence AF210741"
<1..>715
/note="D1 and D2 domains; similar to Heterosigma akashiwo
sequence presented in Genbank Accession Number AF210741"
/product="large subunit ribosomal RNA"
BASE COUNT 166 a 148 c 212 g 189 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 715;
Best Local Similarity 100.0%; Pred. No. 0.15;

QY 1 ACCACGACTGACGACGACCTTT 23
|||||
Db 149 ACCACGACTGACGACGACCTTT 127

RESULT 4
AF211256/c 715 bp DNA linear PLN 11-JAN-2001
LOCUS
DEFINITION Heterosigma sp. CAMR10 large subunit ribosomal RNA gene, partial
sequence.
AF211256
VERSION AF211256.1 GI:12082493
KEYWORDS
SOURCE Heterosigma sp. CAMR10.
ORGANISM Heterosigma sp. CAMR10
Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
REFERENCE 1 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Schollin, C.A., Bergquist, P.R., and Bergquist, P.L.
TITLE Detection and Enumeration of Heterosigma akashiwo and Fibrocapsa
japonica (Raphidophyceae) Using rRNA-targeted Oligonucleotide
Sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Phylogeny of the Raphidophytes Based on Large-Subunit rRNA Gene
Sequences
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 715)
AUTHORS Tyrrell, J.V., Bergquist, P.R., Mackenzie, L., and Bergquist, P.L.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1999) Research and Development, Monterey Bay
Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA
95039-0628, USA
FEATURES
source location/Qualifiers
1..715
/organism="Heterosigma sp. CAMR10"
/strain="CAMR10"
/db_xref="taxon:147349"
/note="similar to Heterosigma akashiwo sequence AF210741"
<1..>715
/note="D1 and D2 domains; similar to Heterosigma akashiwo
sequence presented in Genbank Accession Number AF210741"
/product="large subunit ribosomal RNA"
BASE COUNT 166 a 148 c 212 g 189 t
ORIGIN

Query Match 100.0%; Score 23; DB 8; Length 715;
Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 23: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 ACCAGACTGACGACGACCTT 23
 |||||||
 Db 149 ACCAGACTGACGACGACCTT 127

RESULT 5
 AF409124/c 4103 bp DNA linear PLN 23-OCT-2001
 LOCUS AF409124
 DEFINITION Heterosigma akashiwo internal transcribed spacer 1, 5'8S ribosomal
 RNA gene, internal transcribed spacer 2, and 28S large subunit
 ribosomal RNA gene, complete sequence.
 ACCESSION AF409124 GI:16326556
 VERSION AF409124 GI:16326556
 KEYWORDS
 SOURCE Heterosigma akashiwo.
 ORGANISM Heterosigma akashiwo
 Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
 REFERENCE 1 (bases 1 to 4103)
 AUTHORS Ben Ali, A., De Baere, R., De Wachter, R. and Van de Peer, Y.
 TITLE Evolutionary relationships among heterokont algae (the autotrophic
 stramenopiles) based on a combined analysis of small and large
 subunit RNA
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 4103)
 AUTHORS Ben Ali, A.
 TITLE Direct Submission
 JOURNAL Submitted (14-AUG-2001) Biochemistry, University of Antwerp,
 Universiteitsplein 1, Wilrijk 2610, Belgium
 FEATURES
 source location/Qualifiers
 1..4103
 /organism="Heterosigma akashiwo"
 /db_xref="taxon:2829"
 1..289
 /product="Internal transcribed spacer 1"
 rRNA 290..447
 /product="5.8S ribosomal RNA"
 misc_RNA 448..650
 /product="Internal transcribed spacer 2"
 rRNA 651..4103
 /product="28S large subunit ribosomal RNA"
 BASE COUNT 1045 a 899 c 1142 g 1017 t
 ORIGIN
 Query Match 100.0%; Score 23; DB 8; Length 4103;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCAGACTGACGACGACCTT 23
 |||||||
 Db 824 ACCAGACTGACGACGACCTT 802

RESULT 6
 AC006352/c 126149 bp DNA linear PRI 30-SEP-2000
 LOCUS AC006352
 DEFINITION Homo sapiens PAC clone RP5-1044J9 from 7p14-p12, complete sequence.
 ACCESSION AC006352
 VERSION AC006352.3 GI:4753268
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 126149)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 2 (bases 1 to 126149)
 REFERENCE Du, F., Wohlmann, P. and McGrane, B.
 AUTHORS The sequence of Homo sapiens PAC clone RP5-1044J9
 TITLE

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 126149)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-JAN-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 4 (bases 1 to 126149)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE 5 (bases 1 to 126149)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On May 5, 1999 this sequence version replaced gi:4337277.

COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@watsn.wustl.edu
 ----- Summary Statistics

 Center project name: H.DJ1044J09

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 This clone was derived from human PAC library RPCI-5, prepared by
 Pieter de Jong and coworkers at the Roswell Park Cancer Institute
 (http://bacpac.med.buffalo.edu) using the method described by
 Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from
 one male donor.
 The clone may be obtained either from Genome Systems, Inc.
 (http://www.genomesystems.com) or Research Genetics, Inc.
 (http://www.resgen.com); or from Pieter de Jong.
 VECTOR: pCYPAC2
 NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP5-1044J9;
 actual end is at 126149 of RP5-1044J9.

FEATURES
 source location/Qualifiers
 1..126149
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7p14-p12"
 /clone="RP5-1044J9"
 /clone_id="RPCI-5"
 misc_feature 323..411

```

misc_feature      /note="similar to EST "29670 (NID:g611768)"
337.411
/note="similar to EST AA569186 (NID:g2342240) nm30d10.s1"
misc_feature      494.795
/note="similar to EST AA569186 (NID:g2342240) nm30d10.s1"
misc_feature      575.1056
/note="similar to EST AA946436 (NID:g3109689) on67a01.s1"
repeat_region     1035.1468
/rpt_family="L1"
repeat_region     1590.1602
/rpt_family="TA)n"
repeat_region     1603.1650
/rpt_family="(CA)n"
repeat_region     1651.1661
/rpt_family="(TA)n"
repeat_region     1798.2689
/rpt_family="L1"
repeat_region     2657.2824
/rpt_family="L1"
repeat_region     6125.0206
/rpt_family="MER4-group"
repeat_region     6601.7454
/rpt_family="MER4-group"
repeat_region     7455.7520
/rpt_family="Alu"
repeat_region     7521.8553
/rpt_family="MER4-group"
repeat_region     8554.8830
/rpt_family="Alu"
repeat_region     8831.11043
/rpt_family="MER4-group"
repeat_region     11044.11411
/rpt_family="MaLR"
repeat_region     11412.11453
/rpt_family="MER4-group"
repeat_region     11455.11517
/rpt_family="MER4-group"
repeat_region     11522.11627
/rpt_family="Retroviral"
repeat_region     11614.11909
/rpt_family="LTR198"
repeat_region     11829.12044
/rpt_family="Retroviral"
repeat_region     12045.12268
/rpt_family="MER4-group"
repeat_region     12284.12682
/rpt_family="MER4-group"
repeat_region     12683.12924
/rpt_family="MaLR"
repeat_region     12921.13193
/rpt_family="MER4-group"
repeat_region     14444.14865
/rpt_family="MER4-group"
repeat_region     14961.15275
/rpt_family="Alu"
repeat_region     15394.15993
/rpt_family="MER4-group"
repeat_region     16088.16116
/rpt_family="Al-rich"
repeat_region     16142.16589
/rpt_family="L1"
misc_feature      17241.17338
/note="similar to Mus musculus EST AA170256 (NID:g1748794)
ms87g10.r1"
misc_feature      17241.17338
/note="similar to Mus musculus EST AA575218 (NID:g2349844)
vhl9a01.r1"
misc_feature      17266.17338
/note="similar to EST R72787 (NID:g846819) yj91h09.r1"
misc_feature      17447.17748
/note="similar to EST R72787 (NID:g846819) yj91h09.r1"
misc_feature      17447.17750
/note="similar to Mus musculus EST AA170256 (NID:g1748794)

```

```

misc_feature      ms87g10.r1"
17447.17751
/note="similar to Mus musculus EST AA575218 (NID:g2349844)
vhl9a01.r1"
misc_feature      17490.17780
/note="similar to EST H73816 (NID:g1046750) ys11b01.r1"
misc_feature      17492.17750
/note="similar to EST AA988730 (NID:g3174301) or79h12.s1"
repeat_region     17978.18004
/rpt_family="Ar-rich"
repeat_region     18694.19007
/rpt_family="MER4-group"
repeat_region     20166.20452
/rpt_family="L1"
repeat_region     20456.20498
/rpt_family="L1"
repeat_region     22114.22213
/rpt_family="L1"
misc_feature      22502.22769
/note="match to EST AA489905 (NID:g2220780) ab05c02.r1"
repeat_region     22573.22704
/rpt_family="L1"
misc_feature      23048.23453
/note="similar to EST N52241 (NID:g1193375) yv46q12.s1"
misc_feature      24524.24597
/note="similar to EST AA300571 (NID:g1953132)"
misc_feature      24682.24840
/note="similar to EST AA300571 (NID:g1953132)"
misc_feature      24682.24831
/note="similar to EST T29670 (NID:g611768)"
repeat_region     24954.25460
/rpt_family="L1"
repeat_region     25461.25486
/rpt_family="(CA)n"
repeat_region     25487.25557
/rpt_family="L1"
repeat_region     25650.26548

Query Match      79.1%; Score 18.2; DB 9; Length 126149;
Best Local Similarity 87.0%; Pred. NO. 69;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 ACCAGCAGCAGCAGCAGCAGCCTT 23
Db 123610 ACCAGCAGCAGCAGCAGCAGCCTT 123588

RESULT 7
AL627202/c      169710 bp      DNA      linear      HTG 07-NOV-2001
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone Rpl1-174P23, *** SEQUENCING IN
PROGRESS ***, 12 unordered pieces.
ACCESSION
AL627202
VERSION
AL627202.4 GI:16519605
KEYWORDS
HTG; HTGS_PHASE1; HTGS_ACTIVEIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites)
AUTHORS
Clark, S.
TITLE
Direct Submission
JOURNAL
Submitted (05-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hunger@sanger.ac.uk
On Oct 30, 2001 this sequence version replaced g1:16501251.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hunger@sanger.ac.uk
----- Project Information
Center project name: bA174P23

```

```

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 165366 bases at least Q40
Consensus quality: 167987 bases at least Q30
Insert size: 168610; sum-of-contigs
Insert size: 172230; 12.6% error; agarose-fp
Quality coverage: 5.03x in Q20 bases; sum-of-contigs quality
coverage: 5.15x in Q20 bases; agarose-fp

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 28829: contig of 28829 bp in length
28830 28929: gap of 100 bp
28930 35872: contig of 6943 bp in length
35873 35972: gap of 100 bp
35973 48765: contig of 12793 bp in length
48766 48865: gap of 100 bp
48866 77651: contig of 28786 bp in length
77652 77751: gap of 100 bp
77752 83328: contig of 5577 bp in length
83329 83428: gap of 100 bp
83429 89376: contig of 5948 bp in length
89377 89476: gap of 100 bp
89477 95053: contig of 5577 bp in length
95054 95153: gap of 100 bp
95154 115825: contig of 20672 bp in length
115826 115925: gap of 100 bp
115926 117990: contig of 2065 bp in length
117991 118090: gap of 100 bp
118091 129051: contig of 10961 bp in length
129052 129151: gap of 100 bp
129152 155347: contig of 26196 bp in length
155348 155447: gap of 100 bp
155448 169710: contig of 14263 bp in length.
Location/Qualifiers

```

FEATURES

```

source
1. 169710
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone_lib="RPC1-11.1"
1. 28829
/note="assembly_fragment:00723
fragment_chain:1"
misc_feature
28930..35872
/note="assembly_fragment:03428
fragment_chain:1"
misc_feature
35973..48765
/note="assembly_fragment:02971
fragment_chain:1"
misc_feature
48866..77651
/note="assembly_fragment:02451
fragment_chain:1"
misc_feature
77752..83328
/note="assembly_fragment:02530
fragment_chain:1"
misc_feature
83429..89376
/note="assembly_fragment:01668
fragment_chain:1"
misc_feature
89477..95053
/note="assembly_fragment:02114
fragment_chain:2"
misc_feature
95154..115825
/note="assembly_fragment:03342

```

```

fragment_chain:2"
115926..117990
/note="assembly_fragment:03173
fragment_chain:2"
misc_feature
118091..129051
/note="assembly_fragment:01101
fragment_chain:2"
misc_feature
129152..155347
/note="assembly_fragment:02969"
155448..169710
/note="assembly_fragment:01462
clone_end:17
vector_side:right"

```

```

BASE COUNT 44171 a 42638 c 39303 g 42488 t 1110 others
ORIGIN

```

```

Query Match 79.1%; Score 18.2; DB 2; Length 169710;
Best Local Similarity 87.0%; Pred. NO. 69;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ACCAGACTGACGACGACTTT 23
||||| ||||| ||||| ||
DB 80661 ACCAGACTGACGACGACCTT 80639

```

```

RESULT 8
AL592114 188930 bp DNA linear HTG 03-JAN-2002
LOCUS Homo sapiens chromosome 1 clone RP11-203710, *** SEQUENCING IN
DEFINITION PROGRAMS ***, 2 unordered pieces.
AL592114
VERSION AL592114.9 GI:18070892
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
AUTHORS Clark,S.
TITLE Direct Submission
JOURNAL Submitted (02-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk clone requests: clonequest@sanger.ac.uk
On Jan 4, 2002 this sequence version replaced gi:15501196.

```

COMMENT

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba203710
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M77815; 1% of reads
Sequencing vector: plasmid; 108752; 98% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 186643 bases at least Q40
Consensus quality: 188754 bases at least Q30
Consensus quality: 188773 bases at least Q20
Insert size: 188830; sum-of-contigs
Insert size: 187990; 6.8% error; agarose-fp
Quality coverage: 11.94x in Q20 bases; sum-of-contigs quality
coverage: 12.06x in Q20 bases; agarose-fp

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 34047: contig of 34047 bp in length

```


RESULT 10
AC007882 214025 bp DNA linear PRI 30-SEP-2000
LOCUS
DEFINITION
AC007882 Homo sapiens BAC clone RP11-499D5 from 7p11.2-q11.2, complete
SEQUENCE
AC007882
AC007882.3 GI:7625220
KEYWORDS
HTG.
SOURCE
human
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 214025)
Tittle Sulston, J.E. and Waterston, R.
JOURNAL
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE
99063792
REFERENCE
AUTHORS
2 (bases 1 to 214025)
Tittle Reltz, L., Stoneking, T. and Andrews, S.
JOURNAL
The sequence of Homo sapiens BAC clone RP11-499D5
Unpublished
REFERENCE
AUTHORS
3 (bases 1 to 214025)
Tittle Waterston, R.H.
JOURNAL
Direct Submission
Submitted (19-JUN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA
REFERENCE
AUTHORS
4 (bases 1 to 214025)
Tittle Waterston, R.H.
JOURNAL
Direct Submission
Submitted (20-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA
REFERENCE
AUTHORS
5 (bases 1 to 214025)
Tittle Waterston, R.H.
JOURNAL
Direct Submission
Submitted (21-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA
REFERENCE
AUTHORS
6 (bases 1 to 214025)
Tittle Waterston, R.H.
JOURNAL
Direct Submission
Submitted (22-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA
REFERENCE
AUTHORS
7 (bases 1 to 214025)
Tittle Waterston, R.
JOURNAL
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108,
On Apr 20, 2000 this sequence version replaced g1:5836167.
COMMENT
Genome Center
Center: Washington University Genome Sequencing Center
Center code: WDGSC
Web Site: <http://genome.wustl.edu/gsc>
Contact: saplense@wustl.edu
Summary Statistics
Center Project name: H_NH0499D05

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-499D5; actual end is at base position 214025 of RP11-499D5.

FEATURES

Location/Qualifiers

```

source
1..214025
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="7"
  /map="7p11.2-q11.2"
  /clone="RP11-499D5"
  /clone_id="RPCT-11"
  /size="189..225"
  /rpt_family="AT_rich"
  repeat_region
  335..369
  /rpt_family="AT_rich"
  884..919
  /rpt_family="AT_rich"
  1147..1473
  /rpt_family="MER2_type"
  1473..1519
  /rpt_family="MER2_type"
  1626..1762
  /rpt_family="CATCT)n"
  1735..1893
  /rpt_family="CATCT)n"
  2010..2063
  /rpt_family="MER2_type"
  2138..2152
  /note="similar to EST AA077372 (NID:91836846)"
  repeat_region
  2213..2264
  /rpt_family="MER2_type"
  2288..2433
  /rpt_family="CATCT)n"
  2290..2312
  /note="similar to EST AA077372 (NID:91836846)"
  repeat_region
  2438..2492
  /rpt_family="MER2_type"
  2631..2687
  /rpt_family="MER2_type"
  2650..2675
  /note="similar to EST AA077372 (NID:91836846)"
  misc_feature
  2699..2739
  /rpt_family="MER2_type"
  2808..2864
  /rpt_family="MER2_type"
  2879..3043
  /rpt_family="CATCT)n"
  3140..3201
  /rpt_family="MER2_type"
  3241..3401
  /rpt_family="CATCT)n"

```

```

repeat_region 3498..3646
  /rpt_family="CATCT)n"
misc_feature 3535..3546
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 3764..3822
  /rpt_family="MER2_type"
misc_feature 3796..3821
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 3823..3931
  /rpt_family="CATCT)n"
misc_feature 3866..3875
  /note="similar to EST AA077372 (NID:91836846)"
misc_feature 3951..3971
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 3959..4218
  /rpt_family="L2"
misc_feature 4044..4072
  /note="similar to EST AA077372 (NID:91836846)"
misc_feature 4109..4126
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 4226..4276
  /rpt_family="MER2_type"
repeat_region 4265..4318
  /rpt_family="Achoobo"
repeat_region 4331..4401
  /rpt_family="L1"
misc_feature 4385..4392
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 4412..4571
  /rpt_family="CATCT)n"
misc_feature 4523..4530
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 4572..4815
  /rpt_family="Achoobo"
repeat_region 4824..4917
  /rpt_family="CATCT)n"
misc_feature 4858..4881
  /note="similar to EST AA077372 (NID:91836846)"
repeat_region 4920..4970
  /rpt_family="MER2_type"
repeat_region 4974..5104
  /rpt_family="CATCT)n"
repeat_region 5116..5219
  /rpt_family="L1"
repeat_region 5354..5416
  /rpt_family="MER2_type"
misc_feature 5602..5935
  /note="similar to EST A1806989 (NID:95393555) wI24h01.x1"
repeat_region 5824..5887
  /rpt_family="MER103"
repeat_region 6368..6520
  /rpt_family="Alu"
misc_feature 6368..6806
  /note="similar to EST A1187426 (NID:93738064) qf31d08.x1"
repeat_region 6886..7074
  /rpt_family="MER1_type"
repeat_region 7184..7418
  /rpt_family="Alu"

```

Query Match: 79.1%; Score 18.2; DB 9; Length 214025;
Best Local Similarity 87.0%; Pred. No. 69;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCAGAGCTGACGACGACCTT 23
Db 129136 ACCAGAGAGAGACGACCTT 129158

RESULT 11
AC007908/c 227856 bp DNA linear HTG 25-APR-2001
LOCUS AC007908 Homo sapiens chromosome 16 clone RP11-499D5, WORKING DRAFT
DEFINITION SEQUENCE, 30 unordered pieces.

AC007908
 VERSION AC007908.3 GI:9094205
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 227856)
 DOE Joint Genome Institute.
 TITLE
 JOURNAL Sequencing of Human Chromosome 16
 REFERENCE
 2 (bases 1 to 227856)
 Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
 Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
 Bryant, J., Tesmer, J., Mehncke, L., Longmire, J., White, S., Tatum, O.,
 Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
 McMurry, K., Han, C. and Deaven, L.
 TITLE
 JOURNAL Direct Submission
 Submitted (24-JUN-1999) Center for Human Genome Studies, DOE Joint
 Genome Institute, Los Alamos National Laboratory, MS M888, Los
 Alamos, NM 87545, USA
 On Jul 13, 2000 this sequence version replaced gi:7211873.
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>
 Project Information
 Center Project Name: 0
 Center clone name: RP11-11_499D5

 Summary Statistics
 Consensus quality: 184734 bases at least Q40
 Consensus quality: 201642 bases at least Q30
 Consensus quality: 209314 bases at least Q20
 Estimated insert size: 218590; agarose-ftp estimation
 Estimated insert size: 224956; sum-of-contigs estimation
 Quality coverage: 4.43 in Q20 bases; agarose-ftp estimation
 Quality coverage: 4.3 in Q20 bases; sum-of-contigs estimation
 NOTE: This is a 'working draft' sequence. It currently
 consists of 30 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 1259: contig of 1259 bp in length
 * 1260 1359: gap of unknown length
 * 1360 2455: contig of 1096 bp in length
 * 2456 2555: gap of unknown length
 * 2556 3691: contig of 1136 bp in length
 * 3692 3791: gap of unknown length
 * 3792 4805: contig of 1014 bp in length
 * 4806 4905: gap of unknown length
 * 4906 5938: contig of 1033 bp in length
 * 5939 6038: gap of unknown length
 * 6039 7051: contig of 1013 bp in length
 * 7052 7151: gap of unknown length
 * 7152 8247: contig of 1095 bp in length
 * 8248 8347: gap of unknown length
 * 8348 9994: contig of 1647 bp in length
 * 9995 10093: gap of unknown length
 * 10094 11375: contig of 1282 bp in length
 * 11376 11476: gap of unknown length
 * 11477 12513: contig of 1038 bp in length
 * 12514 12613: gap of unknown length
 * 12614 13645: contig of 1032 bp in length
 * 13646 13745: gap of unknown length
 * 13746 15252: contig of 1507 bp in length
 * 15253 15353: gap of unknown length
 * 15354 16573: contig of 1221 bp in length
 * 16574 16673: gap of unknown length

16674 18205: contig of 1532 bp in length
 * 18206 18305: gap of unknown length
 * 18306 19913: contig of 1608 bp in length
 * 19914 20013: gap of unknown length
 * 20014 22479: contig of 2466 bp in length
 * 22480 22579: gap of unknown length
 * 22580 25293: contig of 2714 bp in length
 * 25294 25394: gap of unknown length
 * 25395 29118: contig of 3725 bp in length
 * 29119 29218: gap of unknown length
 * 29219 32189: contig of 2971 bp in length
 * 32190 32289: gap of unknown length
 * 32290 33294: contig of 1005 bp in length
 * 33295 33394: gap of unknown length
 * 33395 34509: contig of 1115 bp in length
 * 34510 34610: gap of unknown length
 * 34611 36147: contig of 1558 bp in length
 * 36148 36247: gap of unknown length
 * 36248 42248: contig of 6001 bp in length
 * 42249 42348: gap of unknown length
 * 42349 52325: contig of 9977 bp in length
 * 52326 52425: gap of unknown length
 * 52426 60479: contig of 8054 bp in length
 * 60480 60579: gap of unknown length
 * 60580 68336: contig of 8057 bp in length
 * 68337 68737: gap of unknown length
 * 68738 83105: contig of 14369 bp in length
 * 83106 83205: gap of unknown length
 * 83206 116897: contig of 33692 bp in length
 * 116898 116997: gap of unknown length
 * 116998 165561: contig of 48564 bp in length
 * 165562 165661: gap of unknown length
 * 165662 227856: contig of 62195 bp in length.
 Location/Qualifiers
 1..227856
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-499D5"
 /clone_1bp="RP11 human BAC library 11"
 BASE COUNT 68439 a 46078 c 44943 g 64859 t 3537 others
 ORIGIN
 Query Match 79.1%; Score 18.2; DB 2; Length 227856;
 Best Local Similarity 87.0%; Pred. No. 69;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ACCAGCAGTGGAGCAGCAGCTTT 23
 ||||| ||| ||||| ||||| |||
 DB 130616 ACCAGCAGCAGCAGCAGCAGCTTT 130594
 RESULT 12
 HUC002041 234542 bp DNA linear PRI 23-NOV-1999
 LOCUS
 DEFINITION Human Chromosome 16 BAC clone C179875K-A-11E1, complete sequence.
 AC002041
 VERSION AC002041.1 GI:2576343
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 234542)
 Loftus, B.J., Kim, J.J., Shedd, V.P., Kalush, F., Brandon, R.,
 Fuhrmann, J., Mason, T., Crosby, M.L., Barstead, M., Cronin, L.,
 Deslattes Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S.,
 Eichler, E.E., Harris, P.C., Venter, J.C. and Adams, M.D.
 TITLE
 JOURNAL Genome duplications and other features in 12 Mb of DNA sequence
 from human chromosome 16p and 16q
 GENOMICS 99425270
 MEDLINE
 PUBMED 10493829

```

REFERENCE
AUTHORS
2 (bases 1 to 234542)
Adams,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J.,
Brandon,R., Kim,U.J., Kertavagse,A.R. and Venter,J.C.
Human Chromosome 16 BAC clone C11967SK-A-17E1
Unpublished
JOURNAL
3 (bases 1 to 234542)
Adams,M.D. and Loftus,B.J.
Direct Submission
REFERENCE
AUTHORS
Submitted (29-APR-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
JOURNAL
4 (bases 1 to 234542)
Adams,M.D.
Direct Submission
REFERENCE
AUTHORS
Submitted (30-OCT-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
JOURNAL
5 (bases 1 to 234542)
Adams,M.D.
Direct Submission
REFERENCE
AUTHORS
Submitted (31-DEC-1997) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
JOURNAL
6 (bases 1 to 234542)
Adams,M.D. and Loftus,B.J.
Direct Submission
REFERENCE
AUTHORS
Submitted (17-MAR-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
JOURNAL
7 (bases 1 to 234542)
Adams,M.D.
Direct Submission
REFERENCE
AUTHORS
Submitted (11-SEP-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Oct 30, 1997 this sequence version replaced gi:2570252.
Address all correspondence to: Mark Adams The Institute for Genomic
Research 9712 Medical Center Dr.,
Rockville, MD 20850,
USA
COMMENT
e-mail address: mdamas@tigr.org. The bac location is on
chromosome BAC clone is located on human chromosome 16p11.2.
The orientation of the sequence is from Sp6 end to T7 end. Genes
were identified by a combination of five methods including: XGRAIL
(available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder
(Phil Green, University of Washington), Genscan (Chris Burge,
http://genome.stanford.edu/~chris/GENSCANW.html) searches of the
complete sequence against a peptide database,
and the human gene index database at TIGR
(http://www.tigr.org/tdb/hg1/hg1.html). A gene with homology to
another protein is annotated as the isolog of that protein. Genes
without peptide homology having spliced EST hits are termed
'unknown protein'. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES
SOURCE
1..234542
1. "Organism=Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p11.2"
/clone="A-17E1"
36147..39310
/gene="A-17E1.9"
/pseudo
59641..87064
/gene="A-17E1.1"
/note="pseudogene"
/pseudo
87762..87893
/note="32826, A002D07, Chr. -, Homo sapiens"
/db_xref="dbSTS:G19948"
complete(128248..129149)
/gene="A-17E1.2"
/note="pseudogene"
/pseudo
132476..132730
/note="42645, SWSS4047, Chr. 7, Homo sapiens"
/db_xref="dbSTS:G28745"
STS
gene
STS

```

STS	136754..136924 /note="42758, D1S1053, Chr. -, Homo sapiens" /db_xref="dbSTS:c28812"
STS	137107..137228 /note="32826, A002D07, Chr. -, Homo sapiens" /db_xref="dbSTS:C19948"
STS	149776..149953 /note="7821, STS1-cSRU-31b6-uA/cSRU-31b6-uZ, Chr. -, Homo sapiens" /db_xref="dbSTS:G02335"
gene	159399..159940 /gene="A-17E1.8"
STS	160622..160753 /note="32826, A002D07, Chr. -, Homo sapiens" /db_xref="dbSTS:C19948"
STS	163896..164027 /note="32826, A002D07, Chr. -, Homo sapiens" /db_xref="dbSTS:C19948"
misc_feature	complement(163812..202126) /note="creatine transporter and CDM gene containing duplication arising from Chx28"
gene	181687..183736 /gene="A-17E1.4"
STS	/pseudo 190412..190532 /note="18040, SHCG-9731, Chr. 16, Homo sapiens"
gene	complement(190561..197183) /gene="A-17E1.5"
gene	/pseudo 206967..207475 /gene="A-17E1.6"
gene	/pseudo 216986..217560 /gene="A-17E1.10"
gene	/pseudo 228452..229030 /gene="A-17E1.11"
BASE COUNT	63242 a 50172 c 50664 g 70414 t 50 others
ORIGIN	
Query Match	79.1% Score 18.2; DB 9; Length 234542;
Best Local Similarity	87.0%; Pred. No. 69;
Matches	20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY	1 ACCACACTGCAGCAGCCACTT 23 Db 81049 ACCAGTACGAGCACGCCCTT 81071
RESULT 13	
LOCUS	AC012495 215359 bp DNA linear PRI 07-NOV-2001
ACCESSION	AC012495 BAC clone RP11-327f6 from 2, complete sequence.
VERSION	AC012495.9 GI:16756348
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 21535)
TITLE	Sullivan,J.E. and Waterston,R.
JOURNAL	Toward a complete human genome sequence
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)
REFERENCES	99063792
AUTHORS	2 (bases 1 to 21535)
TITLE	Du,F., Haakenson,W., Elliott,G. and Richards,M
JOURNAL	The sequence of Homo sapiens BAC clone RP11-327F6
REFERENCE	(bases 1 to 215359)

AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-OCT-1999) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 215359)

AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 7, 2001 this sequence version replaced gi:11431146.

COMMENT

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 ----- Center project name: H_NH0327F06

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Teleno,M., Cataneese,J.J. and de Jong,P.P. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBac3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-434B12; the clone sequenced to the right is AC051646. Actual start of this clone is at base position 1 of RP11-327F6; actual end is at base position 215359 of RP11-327F6.

FEATURES

source

The sequence H_NH0327F06 from base positions 115714 to 115728 is represented by sequence derived from PCR on BAC DNA.

Location/Qualifiers

1. 215359

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-327F6"

/clone.lib="RPCI-11"

164. .293

/rpt_family="Alu"

repeat_region

473. .757

/rpt_family="Alu"

repeat_region

846. .1142

/rpt_family="Alu"

repeat_region

2744. .3066

/rpt_family="MER2_type"

repeat_region 4273. .4576
 /rpt_family="Alu"
 repeat_region 8207. .8227
 /rpt_family="AT-rich"
 repeat_region 8476. .8607
 /rpt_family="Alu"
 repeat_region 10228. .10586
 /rpt_family="ERV1"
 repeat_region 10727. .10757
 /rpt_family="AT-rich"
 repeat_region 11462. .11768
 /rpt_family="Alu"
 repeat_region 14106. .14201
 /rpt_family="L1"
 repeat_region 14307. .15160
 /rpt_family="L1"
 repeat_region 17498. .17825
 /rpt_family="Alu"
 repeat_region 18341. .18600
 /rpt_family="Alu"
 repeat_region 18942. .19033
 /rpt_family="L2"
 repeat_region 19700. .19759
 /rpt_family="MIR"
 repeat_region 20267. .20410
 /rpt_family="(TTATA)n"
 repeat_region 20576. .20600
 /rpt_family="(A)n"
 repeat_region 20601. .20657
 /rpt_family="(TA)n"
 repeat_region 21725. .22018
 /rpt_family="Alu"
 repeat_region 22203. .22249
 /rpt_family="Alu"
 repeat_region 22316. .22616
 /rpt_family="Alu"
 repeat_region 22657. .22701
 /rpt_family="(T)n"
 repeat_region 22731. .22752
 /rpt_family="(TAA)n"
 repeat_region 24055. .24075
 /rpt_family="AT-rich"
 repeat_region 24507. .24559
 /rpt_family="MIR"
 repeat_region 27464. .27498
 /rpt_family="(T)n"
 repeat_region 27985. .28215
 /rpt_family="Alu"
 repeat_region 28243. .28526
 /rpt_family="Alu"
 repeat_region 28529. .28552
 /rpt_family="AT-rich"
 repeat_region 28752. .28843
 /rpt_family="MIR"
 repeat_region 29211. .29265
 /rpt_family="Mariner"
 repeat_region 29352. .29615
 /rpt_family="L1"
 repeat_region 29647. .29795
 /rpt_family="L1"
 repeat_region 30200. .30220
 /rpt_family="AT-rich"
 repeat_region 31406. .31724
 /rpt_family="L2"
 repeat_region 31839. .32706
 /rpt_family="L1"
 repeat_region 32707. .33079
 /rpt_family="MIR"
 repeat_region 33080. .33160
 /rpt_family="L1"
 repeat_region 34069. .34373
 /rpt_family="Alu"
 repeat_region 34685. .34962

```

repeat_region      /rpl_family="L1"
                    36033..36204
                    /rpl_family="MIR"
repeat_region      /rpl_family="MIR"
                    38668..38979
                    /rpl_family="Alu"
repeat_region      /rpl_family="Alu"
                    39472..39493
                    /rpl_family="MIR"
repeat_region      /rpl_family="MIR"
                    39494..39702
                    /rpl_family="MER2_type"
repeat_region      39703..39785
                    /rpl_family="MIR"
repeat_region      39794..39843
                    /rpl_family="L2"
repeat_region      40006..40612
                    /rpl_family="L1"
repeat_region      40797..40871
                    /rpl_family="L2"
repeat_region      41239..41319
                    /rpl_family="L2"
repeat_region      41909..42079
                    /rpl_family="Alu"
repeat_region      42235..42861
                    /rpl_family="L2"
repeat_region      43465..43500
                    /rpl_family="AT_rich"
repeat_region      45073..45094
                    /rpl_family="AT_rich"
repeat_region      46334..46624
                    /rpl_family="Alu"
repeat_region      46625..46654
                    /rpl_family="AT_rich"
repeat_region      47891..47941
                    /rpl_family="(TA)n"
repeat_region      48406..48703
                    /rpl_family="Alu"

Query Match      77.4% Score 17.8; DB 9; Length 215359;
Best Local Similarity 90.5%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCACGACTGAGCAGCAGCCT 21
    |||||
Db 141309 ACCACGACTGAGCAGCCT 141329

RESULT 14
CNS00YVE      34105 bp      DNA      linear      PRI 11-OCT-1999
DEFINITION    Homo sapiens genomic region containing hypervariable minisatellites
ACCESSION     AL096807.1 GI:5441561
KEYWORDS      hypervariable; minisatellite.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     Petit,E., Vico,V., Brothier,P. and Vergnaud,G.
AUTHORS       1 (bases 1 to 34105)
TITLE         Direct Submission
JOURNAL       Submitted (05-JUL-1999) Laboratoire IECH - Institut de Genetique et
              Microbiologie Universite Paris XI 91400 ORSAY - FRANCE & Genoscope
              - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE
              & UGM - Centre d'Etudes du Bouchet BP3 91710 Verri le Petit - FRANCE
              2 (bases 1 to 34105)
              Genoscope.
REFERENCE     Direct Submission
AUTHORS       Submitted (05-JUL-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
              Location/Qualifiers
              1..34105
              /organism="Homo sapiens"

```

```

/dp_xref="taxon:9606"
/chromosome="8"
/clone="CEBA2"
/map="gq24.3"

BASE COUNT      8661 a 8558 c 8659 g 8227 t

ORIGIN
Query Match      75.7% Score 17.4; DB 9; Length 34105;
Best Local Similarity 94.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCACGACTGAGCAGCAGCACC 20
    |||||
Db 352 CCACGCTGAGCAGCAGCACC 370

RESULT 15
AC021605      197764 bp      DNA      linear      HTG 11-AUG-2001
LOCUS         Homo sapiens clone RP11-225D16, WORKING DRAFT SEQUENCE, 32
DEFINITION    unordered pieces.
ACCESSION     AC021605
VERSION       AC021605.4 GI:15148279
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
AUTHORS       1 (bases 1 to 197764)
TITLE         Homo sapiens, clone RP11-225D16
JOURNAL       Unpublished
AUTHORS       2 (bases 1 to 197764)
              Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barna,N., Beckwith,R., Bede,F.,
              Boguslavsky,L., Bouckhalter,B., Brown,A., Burtell,G., Castle,A.,
              Choquel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
              Dearlano,K., Dewar,K., Domino,M., Doyle,M., Fensholt,J.,
              Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
              Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
              Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
              Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
              McDonald,P., Marquis,N., McKean,P., Mckurk,A., McKernan,K.,
              Mcpheeters,R., Meldrum,J., Menzies,L., Morrow,J., Naylor,J.,
              Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
              Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rotman,D.,
              Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
              Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
              Zimmer,A. and Zody,M.
              Direct Submission
              Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Aug 11, 2001 this sequence version replaced 91:7767808.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence.submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: L5810
              Center clone name: 225-D-16
              ----- Summary Statistics
              Sequencing vector: M13; M77815; 49% of reads
              Sequencing vector: Plasmid; n/a; 51% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 184156 bases at least Q40
              Consensus quality: 190357 bases at least Q30
              Consensus quality: 193088 bases at least Q20

```

Insert size: 170000; agarose-tp
Insert size: 194664; sum-of-ctnigs
Quality coverage: 12.6 in Q20 bases; agarose-tp
Quality coverage: 11.0 in Q20.
NOTE: This is a 'working draft' sequence. It currently
consists of 32 ctnigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the ctnigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 13523: contig of 13523 bp in length
* 13524 13623: gap of 100 bp
* 13624 14303: contig of 680 bp in length
* 14304 14403: gap of 100 bp
* 14404 15055: contig of 652 bp in length
* 15056 15155: gap of 100 bp
* 15156 15778: contig of 623 bp in length
* 15779 15878: gap of 100 bp
* 15879 16229: contig of 351 bp in length
* 16230 16329: gap of 100 bp
* 16330 16887: contig of 558 bp in length
* 16888 16987: gap of 100 bp
* 16988 17233: contig of 246 bp in length
* 17234 17333: gap of 100 bp
* 17334 17999: contig of 666 bp in length
* 18000 18099: gap of 100 bp
* 18100 18777: contig of 678 bp in length
* 18778 18877: gap of 100 bp
* 18878 19460: contig of 583 bp in length
* 19461 19560: gap of 100 bp
* 19561 20742: contig of 1182 bp in length
* 20743 20842: gap of 100 bp
* 20843 21533: contig of 691 bp in length
* 21534 21633: gap of 100 bp
* 21634 22342: contig of 709 bp in length
* 22343 22442: gap of 100 bp
* 22443 23111: contig of 669 bp in length
* 23112 23211: gap of 100 bp
* 23212 23892: contig of 681 bp in length
* 23893 23992: gap of 100 bp
* 23993 24854: contig of 862 bp in length
* 24855 24954: gap of 100 bp
* 24955 26017: contig of 1063 bp in length
* 26018 26117: gap of 100 bp
* 26118 26328: contig of 211 bp in length
* 26329 26428: gap of 100 bp
* 26429 27272: contig of 844 bp in length
* 27273 27372: gap of 100 bp
* 27373 28285: contig of 913 bp in length
* 28286 28385: gap of 100 bp
* 28386 28255: contig of 868 bp in length
* 29254 29353: gap of 100 bp
* 29354 30221: contig of 868 bp in length
* 30222 30321: gap of 100 bp
* 30322 31453: contig of 1132 bp in length
* 31454 31553: gap of 100 bp
* 31554 33045: contig of 1492 bp in length
* 33046 33145: gap of 100 bp
* 33146 33947: contig of 802 bp in length
* 33948 34047: gap of 100 bp
* 34048 35747: contig of 1700 bp in length
* 35748 35847: gap of 100 bp
* 35848 36927: contig of 1060 bp in length
* 36928 37027: gap of 100 bp
* 37028 42128: contig of 5101 bp in length
* 42129 42228: gap of 100 bp
* 42229 50923: contig of 8695 bp in length
* 50924 51023: gap of 100 bp
* 51024 76863: contig of 25840 bp in length
* 76864 76963: gap of 100 bp
* 76964 123889: contig of 46926 bp in length

FEATURES * 123890 123989: gap of 100 bp
* 123990 197764: contig of 73775 bp in length.
source
1. 197764
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-225D16"
/clone_id="RP11-11 Human Male BAC"
1. 13523
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
13624..14303
/note="assembly_fragment"
14404..15055
/note="assembly_fragment"
15156..15778
/note="assembly_fragment"
15879..16229
/note="assembly_fragment"
16330..16887
/note="assembly_fragment"
16988..17233
/note="assembly_fragment"
17334..17999
/note="assembly_fragment"
18100..18777
/note="assembly_fragment"
18878..19460
/note="assembly_fragment"
19561..20742
/note="assembly_fragment"
20843..21533
/note="assembly_fragment"
21634..22342
/note="assembly_fragment"
22443..23111
/note="assembly_fragment"
23212..23892
/note="assembly_fragment"
23993..24854
/note="assembly_fragment"
24955..26017
/note="assembly_fragment"
26118..26328
/note="assembly_fragment"
26429..27272
/note="assembly_fragment"
27373..28285
/note="assembly_fragment"
28386..29253
/note="assembly_fragment"
29354..30221
/note="assembly_fragment"
30322..31453
/note="assembly_fragment"
31554..33045
/note="assembly_fragment"
33146..33947
/note="assembly_fragment"
34048..35747
/note="assembly_fragment"
35848..36927
/note="assembly_fragment"
37028..42128
/note="assembly_fragment"
42229..50923
/note="assembly_fragment"
51024..76863
/note="assembly_fragment"
76964..123889
/note="assembly_fragment"

Query Match 75.7%; Score 17.4; DB 2; Length 197764;
 Best Local Similarity 94.7%; Pred. No. 1.9e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 CCAGCACTGAGCAGCAGC 20
 DB 84705 CCAGCTGTGAGCAGCAGC 84723

Search completed: October 17, 2002, 09:32:13
 Job time : 1466.27 secs